

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»**

**Факультет інформатики та обчислювальної техніки
Кафедра автоматики та управління в технічних системах**

До захисту допущено:

Завідувач кафедри

_____ Олександр РОЛІК

«12» травня 2020 р.

Дипломний проєкт

на здобуття ступеня бакалавра

**за освітньо-професійною програмою «Комп'ютеризовані системи управління»
спеціальності 151 «Автоматизація та комп'ютерно-інтегровані технології»
на тему: «Система автоматизації фінансової звітності та аналітики»**

Виконав (-ла):

студент (-ка) IV курсу, групи ІА-361

Галагуз Валентина Ігорівна

Керівник:

Асистент

Шинкевич Микола Костянтинович

Рецензент:

Посада, науковий ступінь, вчене звання,

Прізвище, ініціали

Засвідчую, що у цьому дипломному
проєкті немає запозичень з праць інших
авторів без відповідних посилань.

Студент (-ка) _____

**Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»**

Факультет інформатики та обчислювальної техніки

Автоматики та управління в технічних системах

Рівень вищої освіти – перший (бакалаврський)

Напрямок підготовки (програма професійного спрямування) – 6. 050201

Системна інженерія

ЗАТВЕРДЖУЮ

Завідувач кафедри

_____ О.І. Ролік

«12» травня 2020 р.

ЗАВДАННЯ

на дипломний проєкт студенту

Галагуз Валентині Ігорівні

1. Тема проєкту «Система автоматизації фінансової звітності засобами Бізнес Аналітики», керівник проєкту Шинкевич Микола Костянтинович, Асистент, затверджені наказом по університету від «12» травня 2020 р. №1089-с

2. Термін подання студентом проєкту _____

3. Вихідні дані до проєкту

4. Зміст пояснювальної записки

5. Перелік графічного матеріалу (із зазначенням обов'язкових креслеників, плакатів, презентацій тощо)

7. Дата видачі завдання 20 грудня 2019 р.

Календарний план

| № з/п | Назва етапів виконання дипломного проєкту | Термін виконання етапів проєкту | Примітка |
|-------|--|---------------------------------|----------|
| 1 | Затвердження теми дипломного проєкту | 20.12.2019 | |
| 2 | Ознайомлення з літературними джерелами | 15.03.2020 | |
| 3 | Аналіз існуючих рішень | 01.04.2020 | |
| 4 | Моделювання компонент розроблюваної системи | 10.04.2020 | |
| 5 | Встановлення компонент системи на підприємстві | 30.04.2020 | |
| 6 | Тестування та прийняття системи змовником | 08.05.2020 | |
| 7 | Створення технічної документації | 15.05.2020 | |
| 8 | Написання пояснювальної записки | 23.05.2020 | |
| 9 | Передзахист проєкту | 02.06.2020 | |
| 10 | Захист проєкту | 17.06.2020 | |

Студентка

Галагуз В.І

Керівник проєкту

Шинкевич М.К

АНОТАЦІЯ

Галагуз В.І. Система автоматизації фінансової звітності та аналітики. КПП ім. Ігоря Сікорського, Київ, 2020.

Проект містить 62 с. тексту, 6 рисунків, 4 таблиці, 6 додатків та 22 джерел, загальним обсягом 67 сторінок.

Ключові слова: система, бізнес аналітика, звітність, Вітрина Даних, інфраструктура

Пояснювальна записка дипломного проекту складається з чотирьох розділів, містить 4 таблиці, 6 рисунків, 6 додатків та 22 джерел – загалом 67 сторінок.

Об'єкт дослідження: система фінансової звітності та аналітики що використовується для побудови звітів та аналізу даних

Мета дипломного проекту: спроектувати та розробити комплекс засобів системної інфраструктури, що забезпечують повноцінне функціонування платформи Бізнес Аналітики та мають засоби зберігання й обробки даних фінансового домену.

У першому розділі було досліджено актуальність поставленої задачі та досліджено декілька підходів до технічної реалізації комплексу засобів системи аналітики.

У другому розділі було розроблено вимоги до системи та поставлено критерії перевірки відповідності системи даним вимогам.

У третьому розділі було проведено дослідження існуючих засобів реалізації трьох рівнів системи та порівняння їх між собою. В результаті було обрано засоби розгортки для компонент системи.

У четвертому розділі наведено алгоритм та деталі розгортки кожного рівня системи. Описано процеси налаштування та взаємо-зв'язки рівнів системи. Було оцінено продуктивність роботи системи.

SUMMARY

Halahuz V. Automation system of finance reporting and analysis. Igor Sikorsky KPI, Kyiv, 2020.

Keywords: system, business analytics, reporting, Data Mart, infrastructure

The explanatory note of the diploma project consists of four sections, contains tablei, 6 applications and 22 sources - a total of 62 pages.

The object of study: finance reporting and analysis system which is designated for creating reports and data analysis

The aim of the diploma project: to model and build complex solution of system infrastructure to support functioning of Business Analysis platform and include tools for financial data processing and storage.

There was a research of relevancy for a stated task conducted together with investigation of a few methods of deployment technical components in the first section.

The second section includes requirements statement to the system and specifies acceptance criterias according to the requirements.

The third section conducts research and comparison of already existing solutions. As the result tools for system layers deployment were chosen.

Forth section represents deployment algorithm and milestones for every part of the system. Processes of setup and cross-relations of different layers and components were described. System performance checks were conducted.

Пояснювальна записка
до дипломного проєкту
на тему: «Система автоматизації фінансової
звітності та аналітики»

Київ – 2020 року

ЗМІСТ

| | |
|---|----|
| ВСТУП | 4 |
| 1. ПОСТАНОВКА ЗАДАЧІ..... | 6 |
| 1.1 Призначення та галузі використання систем Бізнес Аналітики . | 6 |
| 1.2 Напрямки розвитку Бізнес аналітики..... | 9 |
| 1.3 Розгляд підходів до збереження та обробки даних | 10 |
| Висновки до розділу | 16 |
| 2. РОЗРОБКА ВИМОГ ТА ХАРАКТЕРИСТИК СИСТЕМИ | 17 |
| Висновки до розділу | 21 |
| 3. АНАЛІЗ І ПРИЙНЯТТЯ РІШЕНЬ ДЛЯ РЕАЛІЗАЦІЇ СИСТЕМИ..... | 22 |
| 3.1 Огляд існуючих платформ Бізнес Аналітики..... | 22 |
| 3.1.1 Microsoft Power BI | 23 |
| 3.1.2 Tableau | 25 |
| 3.1.3 Qlik | 26 |
| 3.2 Реляційні бази даних | 28 |
| 3.2.1 База даних SQL Server 2017..... | 28 |
| 3.2.2 Amazon Aurora..... | 31 |
| 3.3 Обчислювальна хмара Amazon Elastic (Amazon EC2) | 34 |
| Висновки до розділу | 42 |
| 4 РЕАЛІЗАЦІЯ ПРОГРАМНО-АПАРАТНИХ ЗАСОБІВ БІЗНЕС АНАЛІТИКИ | 44 |
| 4.1 Потоки даних | 44 |
| 4.2 Затвердження алгоритму реалізації системи..... | 45 |
| 4.2.1 Реалізація системного рівня..... | 45 |

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 2 |
| Зм | Лист | № документа | Підпис | Дата | | |

| | | |
|-------|--|----|
| 4.2.2 | Налаштування Баз даних для Вітрини даних | 49 |
| 4.2.3 | Розгортання платформи Бізнес Аналітики..... | 50 |
| 4.3 | Розробка Вітрини Даних..... | 55 |
| 4.4 | Результати реалізації системи | 57 |
| | Висновки до розділу | 59 |
| | ВИСНОВКИ..... | 60 |
| | СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ..... | 62 |

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 3 |
| Зм | Лист | № документа | Підпис | Дата | | |

ВСТУП

Планування виробництва, підрахунки бюджету та прогнозування в сьогоdnішній швидко ростучій та змінюваній глобальній економіці вимагає від корпоративних установ бути гнучкими та здатними швидко реагувати на нові ринкові можливості та загрози. Успіх підприємств вимагає наявності засобів для постановки та контролю за досягненням корпоративних цілей та бюджетів у поєднанні з можливістю постійно відслідковувати продуктивність та своєчасно обновляти заплановані цільові показники планування, розподілу ресурсів використовуючи сучасні засоби.

Бізнес Аналітика, це термін що в даний час інстинктивно асоційований з інформаційними технологіями та почав свій розвиток близько 150 років тому. Хоч початок розвитку Бізнес Аналізу відбувся раніше ніж винайдення та поширення комп'ютерних технологій, активний час розвитку цієї сфери почався відповідно до еволюції комп'ютерів та баз даних.

Аналітичні застосунки в банківській та фінансовій індустрії допомагають систематизувати та аналізувати як внутрішні так і зовнішні дані, що знаходить своє відображення в розумних підходах до засобів управління бізнесом та прийняттю кращих управлінських рішень.

Управлінський підхід оснований на такому підході виявляє свою продуктивність в приношенні прибутків, зменшенню ризиків та створенню конкурентної переваги.

З розвитком інформаційних систем кількість даних для аналізу та побудови звітів невинно збільшується. Наразі масштаби оброблюваної інформації неможливо опрацювати навіть з 50% відсотком автоматизації процесу. Цінністю даного проекту є можливість повної автоматизації процесу фінансової звітності, що покращує якість оброблюваних даних та зменшує ризик отримання похибки в обрахунках. Також зменшення затрат людино-годин для аналізу та побудови звітів допомагає скоротити витрати компаній на

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 4 |
| Зм | Лист | № документа | Підпис | Дата | | |

утримання аналітиків та направити вже наявні ресурси на рівень підтримки прийняття рішень.

Метою та завданням даного проекту є проектування та розгортка цілого комплексу засобів для функціонування та підтримки Аналітичної платформи.

В ході реалізації дипломного проекту реалізовано всі рівні системи автоматизації фінансової звітності та проведено тестування з ключовою групою користувачів. Оцінено ефективність роботи системи в кількості зменшених робочих годин на виконання цього завдання з використанням засобів часткової автоматизації процесу.

Пояснювальна записка складається з чотирьох розділів. Графічна частина включає п'ять креслеників формату А3

Загальний обсяг проекту складає 62 сторінки.

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 5 |
| Зм | Лист | № документа | Підпис | Дата | | |

1. ПОСТАНОВКА ЗАДАЧІ

1.1 Призначення та галузі використання систем Бізнес Аналітики

Інструменти, які використовуються для обробки даних клієнтів для отримання висновків можуть призвести до розумнішої практики управління та кращих бізнес-рішень. З цією метою використовуються рішення Business Intelligence (BI) для підвищення прибутковості, зниження ризику та створення конкурентних переваг.

Підвищення операційної ефективності – на сьогоднішньому конкурентному ринку корпоративні установи повинні бути максимально гнучкими та ефективними. Використовуючи рішення BI для аналізу операційних процесів, організації можуть знизити поточні витрати та максимізувати наявні ресурси та досвід. Наприклад, проаналізувавши ефективність роботи співробітників, що звертаються до клієнтів, таких як торговий персонал, продавці та менеджери облікових записів, організації можуть виявити шляхи вдосконалення та покращення досвіду клієнтів у точці контакту.

Покращені продукти та послуги – рішення BI дозволяють організаціям відстежувати окремі потоки доходів, щоб краще визначити, які продукти та послуги вигідні, а які ні. Але переваги на цьому не зупиняються.

Рішення Business Intelligence також дозволяють організаціям аналізувати величезну кількість даних про клієнтів, щоб отримати уявлення про потреби та настрої клієнтів, які можна використовувати для вдосконалення продуктів та послуг. Як приклад, клієнти хочуть швидшого, простішого способу відстеження та аналізу статусу й кількості замовлень в інтернет-магазині. Установи, можливо, зможуть надсилати клієнтам більш своєчасні сповіщення про статус доставки, надавати прості засоби візуалізації аналізу витрат та актуальні поради для повторного замовлення товарів побутового вжитку. Озброївшись

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 6 |
| Зм | Лист | № документа | Підпис | Дата | | |

подібними відомостями, організації можуть розробляти нові та вдосконалені програмні продукти та послуги, розширяти асортимент вітрини замовлень для кращого задоволення потреб клієнтів, а в свою чергу отримують конкурентну перевагу. [1]

Покращений маркетинг – використовуючи ВІ, маркетологи можуть аналізувати дані CRM на основі цілого ряду критеріїв, щоб виявити найвигідніший профіль клієнта. Крім того, клієнтську базу можна проаналізувати для виявлення та розробки нових можливостей перехресного продажу та розпродажів, а також для проведення більш цілеспрямованих онлайн-маркетингових кампаній. Це є головною перевагою, оскільки дослідження показують, що продавати продукти та послуги новим клієнтам коштує в п'ять разів більше, ніж існуючим клієнтам.

Покращене утримання клієнтів – програми ВІ можуть допомогти фінансовим установам визначити та утримувати тих клієнтів, які є найвигіднішими. ВІ також відіграє важливу роль у покращенні утримання клієнтів та лояльності. Використовуючи інструменти та методи бізнес-аналітики, організації можуть виявити причини, через які клієнти переходять на конкуруючу установу. Потім вони можуть впроваджувати нові процеси, щоб допомогти зменшити тиск клієнтів. Можливість відстежувати звички, уподобання та поведінку клієнтів також дозволяє організаціям адаптувати свої продукти та послуги таким чином, щоб вони відповідали потребам, вирішували проблеми та сприяли утриманню та лояльності клієнтів.

Розробка нових інвестиційних стратегій – керівники активів використовують нові набори даних для розробки нових стратегій інвестування. Розробляючи моделі навколо соціальних медіа, інвестори можуть отримати конкретне уявлення про настрої та розвинути торгові сигнали. Інші аналітики досліджень використовують супутникові знімки, щоб зрозуміти глобальну пропозицію товарів, таких як нафта та газ, або триангуляцію споживчих витрат, виходячи з кількості автомобілів на стоянках торгових центрів. Цілі нові

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 7 |
| Зм | Лист | № документа | Підпис | Дата | | |

категорії інвестицій з'являються завдяки використанню аналітики та ВІ-додатків.

Зменшення ризику – фінансовий світ постійно змінюється і наповнюється невизначеністю. Бажано, ніж будь-коли, банківським та фінансовим установам необхідно використовувати всі наявні інструменти для зменшення ризику. Сьогоднішні рішення Business Intelligence надають корисну інформацію, яку організації можуть використовувати для зменшення ризику в ряді областей.

Можливість відстежувати історію транзакцій клієнтів дозволяє установам швидко виявляти та зменшувати випадки шахрайських дій, найбільш помітними є шахрайство з кредитними картками. Можливість відстежувати комунікації та поведінку внутрішніх службовців при торгівлі цінними паперами допомагає установам виконувати нові рамки регулювання, спричинені фінансовою кризою 2008 року та випадками торгівлі інсайдерами. Розблокування даних із систем класів активів може допомогти світовим банкам передбачити кредитний ризик для контрагентів у всіх класах активів.

В майбутньому ті установи, які приймають і повністю використовують рішення ВІ для управління ризиками, підвищення операційної ефективності та надання продуктів і послуг, що відповідають реальним потребам клієнтів, будуть краще позиціонуватись для постійного зростання, прибутковості та конкурентної переваги на довгі роки.

Стовідсоткова автоматизація звітності засобами Бізнес Аналітики налічує рівні:

- інтерфейс застосунку (аналітична платформа, що має можливості побудови, перегляду та зміни звітів);
- моделювання (створення програмного засобу або процесу, що опрацьовує вхідні дані для подальшого використання засобами візуалізації);
- рівень даних (проектування та налаштування сховища даних);

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 8 |
| Зм | Лист | № документа | Підпис | Дата | | |

- технічна інфраструктура (встановлення та налаштування програмно-технічного комплексу засобів для розміщення та підтримки програмних компонент системи в цілому).

1.2 Напрямки розвитку Бізнес аналітики

В даний час ВІ в цілому стає на перше місце з темпами розвитку посеред інших інформаційних технологій. Саме тому через широкий спектр використання почали виділяти декілька різних типів платформ та напрямків бізнес аналітики.

Штучний інтелект (Artificial Intelligence) – звіт про тенденції технологій Gartner ставить AI на 1 місце. Штучний інтелект та машинне навчання тепер виконують складні завдання, які зазвичай виконуються інтелектом людини. Ця можливість використовується для створення аналізу даних у режимі реального часу та звітування на інформаційній панелі, щоб побачити ситуації, як вони відбуваються щосекунди, коли машини виконують завдання самостійно. [9]

Колаборативний ВІ – це програмне забезпечення для бізнес-аналітики, поєднане з інструментами спільної роботи, включаючи соціальні медіа та інші новітні технології для покращення спільної роботи та обміну даними та інформацією командами для спільного прийняття рішень. Термін "ВІ із самообслуговуванням" потрапляє до цієї категорії, оскільки охоплює засоби ВІ, які можуть використовувати будь-який хтось поза командою ІТ для доступу та інтерпретації даних.

Вбудований ВІ – це включає інтеграцію програмного забезпечення ВІ або деяких його функцій в інший бізнес-додаток, щоб розширити його аналітичну чи функціональну звітність. Можливості, звичайно характерні для програмного забезпечення ВІ, вбудовуються в інший не ВІ додаток, тим самим впорядковуючи збір та аналіз даних. Це породить середовище "аналітика скрізь".

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 9 |
| Зм | Лист | № документа | Підпис | Дата | | |

Хмарні обчислення (Cloud Analytics) – Через більш доступну вартість та скорочення часу розгортання в хмарі буде запропоновано більше ВІ-додатків, і більше підприємств перейде на цю технологію.

Враховуючи приведені описи можна визначити, що для згідно із задачею проекту для реалізації необхідно дослідити наявні засоби колаборативної аналітики.

Важливим є метод розгортки рішення. Хмарний та локальний методи порівняно в таблиці 1.1.

Таблиця 1.1. – Порівняння характеристик Хмарних та Локальних рішень

| Характеристика | Хмарний | Локальний |
|------------------------------------|--|--------------------------------------|
| час впровадження | швидкий | в цілому значно довше |
| попередні витрати | низькі | Високі |
| додаткові витрати на устаткування | немає | Є |
| рівень персоналізації | зазвичай низький | значні можливості до персоналізації |
| контроль рівня безпеки даних | на рівні постачальника рішення | цілком управляється організацією |
| довгострокові витрати на утримання | передбачувані (ліцензування, оренда сховища, тощо) | непередбачувані (але ймовірно нижчі) |

1.3 Розгляд підходів до збереження та обробки даних

Сховище даних (Data Warehouse) – це центральне сховище інформації, які подаються для подальшого аналізу для прийняття більш обґрунтованих рішень. Дані надходять у сховище даних з транзакційних систем, реляційних баз даних

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 10 |
| Зм | Лист | № документа | Підпис | Дата | | |

та інших джерел, як правило, на регулярній основі. Бізнес-аналітики, аналітики даних та особи, які приймають рішення, отримують доступ до даних за допомогою інструментів бізнес-аналітики (BI), клієнтів SQL та інших програм для аналітики. [17]

Компанії використовують звіти, інформаційні панелі та інструменти аналітики для отримання інформації про свої дані, моніторингу ефективності бізнесу та підтримки прийняття рішень. Ці звіти, інформаційні панелі та інструменти аналітики працюють на базі сховищ даних, які ефективно зберігають дані для мінімізації вводу-виводу та надання результатів запитів із високою швидкістю сотням і тисячам користувачів одночасно.

Сховища даних, як правило, використовуються для кореляції різноманітних бізнес-даних для забезпечення кращого розуміння керівництвом діяльності компанії.

Багато типів бізнес-даних аналізуються через сховища даних. Потреба в сховищі даних часто стає очевидною, коли аналітичні вимоги не відповідають поточному виконанню операційних баз даних. Запуск складного запиту в базі даних вимагає від бази даних ввести тимчасовий рівень ізоляції для виконання операцій. Це часто не під силу для транзакційних баз даних. Сховище даних використовується для здійснення аналітичної роботи, залишаючи транзакційну базу даних вільною, щоб зосередитися на обробці транзакцій.

Інші переваги сховища даних – це можливість аналізу даних з різних джерел та узгодження відмінностей у схемі зберігання даних за допомогою процесу ETL (Extract Transform Load).

Архітектура сховища даних складається з трьох рівнів. Нижній рівень архітектури – це сервер баз даних, де завантажуються та зберігаються дані. Середній рівень складається з механізму аналітики, який використовується для доступу та аналізу даних. Верхній рівень – це клієнт, що представляє результати за допомогою інструментів звітування, аналізу та пошуку даних.

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 11 |
| Зм | Лист | № документа | Підпис | Дата | | |

Сховище даних працює, організовуючи дані в схему, яка описує макет і тип даних, таких як ціле число, поле даних або рядок. Коли дані приймаються, вони зберігаються в різних таблицях, описаних схемою. Інструменти запитів використовують схему для визначення таблиць даних для доступу та аналізу.

Переваги, що надає використання сховища даних:

- краще прийняття рішень;
- консолідація даних з багатьох джерел;
- якість даних, послідовність та точність даних;
- зберігання історичних даних;
- відокремлює обробку аналітики від транзакційних баз даних, покращуючи продуктивність обох систем.

Сховище даних та бази даних. Сховище даних спеціально розроблений для аналітики даних, що передбачає зчитування великої кількості даних для розуміння взаємозв'язків та тенденцій даних. База даних використовується для збору та зберігання даних, таких як запис деталей транзакції. [7]

В таблиці 1.2. наведено порівняльну характеристику Сховища даних та транзакційної бази даних.

Таблиця 1.2. – Порівняльна характеристика Сховища даних та транзакційної Бази Даних

| характеристика | Сховище даних | Транзакційна База даних |
|-------------------------|---|---|
| відповідні навантаження | аналітика, звітність, обробка даних великими даними | обробка транзакцій |
| джерело даних | дані, зібрані та нормалізовані з багатьох джерел | дані, зібрані як є, є з одного джерела, наприклад транзакційної системи |

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 12 |
| Зм | Лист | № документа | Підпис | Дата | | |

Продовження таблиці 1.2. – Порівняльна характеристика Сховища даних та Транзакційної Бази Даних

| Характеристика | Сховище даних | Транзакційна База даних |
|----------------------|--|---|
| операції збору даних | операції масового запису, як правило, за заздалегідь визначеним пакетним графіком | оптимізований для операцій безперервного запису, оскільки доступні нові дані для максимальної пропускної здатності транзакцій |
| нормалізація даних | денормалізовані схеми, такі як схема зірка або сніжинка | високо нормалізовані статичні схеми |
| зберігання даних | оптимізовано для простоти доступу та високошвидкісного виконання запитів за допомогою колонарного зберігання | оптимізовано для високоефективних операцій запису до одного фізичного блоку, орієнтованого на рядки |
| доступ до даних | оптимізовано для мінімізації вводу-виводу та максимальної пропускної здатності даних | великі обсяги малих операцій зчитування |

Сховище даних та Озеро даних (Data lake). На відміну від сховища даних, озеро даних є централізованим сховищем для всіх даних, включаючи структуровані та неструктуровані. У сховищі даних використовується заздалегідь визначена схема, оптимізована для аналітики. У озері даних схема не визначена, що дозволяє¹ використовувати такі типи аналітики, як аналітика великих даних, пошук повнотекстового тексту, аналітика в режимі реального часу та машинне навчання. [18]

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 13 |
| Зм | Лист | № документа | Підпис | Дата | | |

Порівняльна характеристика Сховища даних із Озером даних наведена в таблиці 1.3.

Таблиця 1.3. – Порівняльна характеристика Сховища даних та Озера даних

| Характеристики | Сховище даних | Озеро даних |
|---------------------|--|---|
| Дані | реляційні дані з транзакційних систем, операційних баз даних та ліній ділових додатків | нереляційні та реляційні дані з пристроїв IoT, веб-сайтів, мобільних додатків, соціальних медіа та корпоративних додатків |
| Схема | розроблена до реалізації сховища даних (схема-на-записі) | написана під час аналізу (схема-на-читанні) |
| ціна / ефективність | найшвидші результати запитів із використанням більш високих витрат на зберігання | результати запиту стають швидшими, використовуючи недороге сховище |
| якість даних | високоякісні дані, які служать центральним "джерелом правди" | будь-які дані, які можуть бути, а можуть і не бути очищеними (тобто необроблені дані) |
| користувачі | бізнес-аналітики, аналітики даних та розробники даних | науковці, розробники даних та бізнес-аналітики (з використанням опрацьованих даних) |
| Аналітика | аналіз пакетної звітності, бі та візуалізації | машинне навчання, прогнозована аналітика, відкриття даних та профілювання |

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 14 |
| Зм | Лист | № документа | Підпис | Дата | | |

Сховище даних та Вітрина даних (Data Mart). Вітрина даних – це сховище даних, яке обслуговує потреби конкретної команди чи бізнес-підрозділу, наприклад, фінанси, маркетинг чи продажі. Він менший, більш зосереджений і може містити резюме даних, які найкраще обслуговують його користувачів. Вітрина даних орієнтована на єдину функціональну область організації та містить підмножину даних, що зберігаються у сховищі даних.

Вітрина даних зазвичай черпає дані лише з кількох джерел порівняно зі сховищем даних. Дані даних мають невеликі розміри і є більш гнучкими порівняно зі Сховищем даних. [18]

Порівняльна характеристика Сховища даних та Вітрини даних наведена в таблиці 1.4.

Таблиця 1.4. – Порівняльна характеристика Сховища даних та Вітрини даних

| характеристики | Сховище даних | Вітрина даних |
|----------------------|--|--|
| область застосування | централізована, декілька предметних областей, об'єднаних разом | децентралізована, конкретна предметна область |
| користувачі | вся компанія | єдине співтовариство чи відділ |
| джерело даних | багато джерел | одне або декілька джерел або частина даних, вже зібраних у сховищі даних |
| розмір | великий, може становити 100 гігабайт до петабайт | малий, як правило, до 10 гігабайт |
| дизайн | зверху вниз | знизу вгору |
| деталізація даних | повні, детальні дані | можуть містити узагальнені дані |

Висновки до розділу

Провівши дослідження предметної області зроблено висновок, що розгортка та встановлення системи Бізнес Аналітики є складним проектом, який налічує декілька рівнів реалізації:

- системний рівень (технічно-серверна складова);
- рівень збереження та обробки даних;
- встановлення та налаштування платформи Бізнес Аналітики.

Розглянуто декілька підходів до збереження даних. Визначено, що найкращим підходом до реалізації рівня даних є підхід Вітрини даних.

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 16 |
| Зм | Лист | № документа | Підпис | Дата | | |

2. РОЗРОБКА ВИМОГ ТА ХАРАКТЕРИСТИК СИСТЕМИ

Функціональні вимоги системи – це підмножина (більш детальний вигляд) бізнес-вимог з точки зору функціональності користувача. Функціональні вимоги визначаються з точки зору бажаної / необхідної взаємодії між кінцевими користувачами бізнесу та цільовим додатком. [6]

Впровадження та побудова системи засобів Бізнес Аналізу являє собою складний проект що нараховує декілька рівнів реалізації. Це зумовлено основною вимогою до таких систем: Швидка реакція на зміни в бізнесі та кількості й структурі даних.

Наведені нижче напрямки вимог до бізнес-аналітики є загальними та необхідними в будь-якій інтеграції ВІ:

Функції платформи: функції програмного забезпечення Бізнес Аналітики встановлюють базову лінію системи. Такі функції, як інтерактивність, простота використання та налаштування, впливають на функціональність інших інструментів програмного забезпечення:

- інтерактивна візуалізація;
- зручний для користувачів;
- налаштування платформи.

Візуалізація даних: важливою вимогою програмного забезпечення для бізнес-аналітики є представлення аналізів на візуальному та доступному носії. Звіти організовують розуміння даних у форматі, який можна легко налаштувати, поділитись та інтерпретувати. Завдяки особливостям, орієнтованим на налаштування, користувачі можуть коригувати звіти, щоб вони виглядали згідно вимогам, та відображати інформацію, якою вони повинні ділитися. Обов'язковою є наявність наступних засобів візуалізації:

- сторінкова організація звітів;
- інтеграція з картами (географічно-просторова);
- анімації;

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 17 |
| Зм | Лист | № документа | Підпис | Дата | | |

- таблиці/матриці;
- діаграми та графіки;
- фільтри;
- можливість перегляду ієрархічних даних;
- шаблони;
- можливість управління даними мовою SQL;

і так далі.

Самообслуговування: користувач має отримати змогу створювати нові звіти та представлення даних без залучення технічного відділу або програмістів. Для цього платформа має бути простою у вивченні, зрозумілою для користувача та використовувати найбільш гнучкі методи управління. При задоволенні цієї вимоги не потрібно проводити навчання для початку та лише мінімальне навчання досвідченого бізнес-користувача.

Аналітика: Аналітичні функції надають уявлення про бізнес-процеси за допомогою оцінки прибутку, продуктивності, соціальних медіа та показників місцеположення. Системи бізнес-аналітики використовують аналіз та прогнозування в реальному часі для спостереження за поточною ефективністю та прогнозуванням майбутніх можливостей:

- бенчмаркінг;
- прогнозована аналітика;
- аналітика соціальних медіа;
- веб-аналітика;
- геолокаційний аналіз;
- показники тенденцій;
- аналіз прибутку;
- статистична аналітика;
- машинне навчання.

Online analytical processing (OLAP): OLAP включає багатовимірний аналіз веб-кубів даних, в яких розміщені великі шаруваті набори інформації.

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 18 |
| Зм | Лист | № документа | Підпис | Дата | | |

Можливості розгортання та дослідження даних забезпечують доступ до деталей цих баз даних. Доступ до цих шарів даних показує, як різні змінні співвідносяться один з одним і розкриває можливі події:

- багатовимірний аналіз;
- ієрархічна навігація;
- дослідження даних;
- автогенерація часових рядів;

Управління документами: можливість експорту звітів у різні формати документів дозволяє спростити обмін інформацією. Користувачі програмного забезпечення ВІ можуть перетворити візуалізацію даних у формат, найбільш засвоюваний для своєї аудиторії – чи то це збирання інформації в електронній таблиці, чи просто перетворення звіту в PDF для простого обміну:

- експорт до Microsoft Excel;
- експорт до робочої книги Microsoft;
- експорт у PDF;
- експорт у HTML;
- версія;

Безпека та захист даних: безпека є однією з найважливіших задач у програмному забезпеченні, особливо при розгортці у хмарних середовищах. Тому параметр рівня доступу користувача є однією з найважливіших вимог для рішення ВІ. Цей параметр дозволяє встановити обмеження щодо того, хто може отримати доступ, експортувати та редагувати ваші файли, електронні таблиці та графіки.

Інтеграції: порівнюючи інструменти ВІ можна виділити окрему групу платформ, які обмежують своє програмне забезпечення власною архітектурою. Однак найкраще програмне забезпечення для ВІ засноване на відкритій архітектурі. Це спрощує інтеграцію з іншими програмними засобами. Інструмент ВІ також повинен дозволяти користувачам об'єднувати дані з різних програм. Наприклад, графіки та діаграми є невід'ємною частиною інструментів

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 19 |
| Зм | Лист | № документа | Підпис | Дата | | |

візуалізації BI і повинні інтегруватися у звіти. Імпорт програми заощаджує кінцевим користувачам час, необхідний для повторного відтворення цих функцій:

- інтеграція ERP;
- інтеграція ETL;
- інтеграція порталів;
- інтеграція CRM;
- програми для MS Office;

та ін.

Інтеграція з різними видами джерел даних та Big Data: найпопулярніші інструменти BI не обмежують їх підтримку однією конкретною базою даних. В ідеалі найкращі програми дозволяють отримати доступ до різноманітних джерел даних. Ці джерела даних повинні включати сховища даних, сховища даних, SQL-сервери тощо. Найпоширенішими є наступні джерела даних:

- Hadoop;
- Hive;
- Hbase;
- Cassandra;
- MapReduce;
- MS SQL server;
- My SQL;
- Redshift;
- PostgreSQL;
- файлові дані;

та ін.

Розгортання: обраний спосіб розгортання залежить від існуючої інфраструктури бізнесу та вимог до сховища даних та безпеки інформації. Встановлення комплексу засобів засобами власних ресурсів інфраструктури задовольняють потреби організації, що має локально зберігати дані.

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 20 |
| Зм | Лист | № документа | Підпис | Дата | | |

Використання засобів хмарних технологій розгортання забезпечує безпеку та постійний доступ при потребі. Мобільні застосунки надають можливість отримати доступ до вашого ВІ-рішення з будь-якого пристрою.

- за вимогою (on-premise);
- хмара;
- локальна установка;
- мобільний.

Будь-яка оцінка нефункціональних або бізнес-аналітичних функціональних вимог до програмного забезпечення ВІ має бути виконана орієнтуючись на задоволення вище наведених критеріїв.

При розробці проекту було враховано визначені вимоги.

Висновки до розділу

В даному розділі було визначено вимоги до розроблюваної системи. Напрямки вимог розглянуто з точки зору функціональних та технічних вимог до розроблюваного продукту.

Дані вимоги використано для визначення системи що найкраще відповідає потребам користувачів та забезпечує стабільність роботи рішення.

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 21 |
| Зм | Лист | № документа | Підпис | Дата | | |

3. АНАЛІЗ І ПРИЙНЯТТЯ РІШЕНЬ ДЛЯ РЕАЛІЗАЦІЇ СИСТЕМИ

3.1 Огляд існуючих платформ Бізнес Аналітики.

Програмне забезпечення Великих даних, аналітики та бізнес-аналітики зазнає рекордного зростання. За даними Forbes, три фактори сприяють розвитку великих даних, аналітики та ринку ВІ:

- перехід підприємств на більш орієнтований на клієнта підхід;
- зусилля для поліпшення операційної діяльності;
- вихід підприємств на нові ринки та прийняття нових бізнес-моделей.

Питання вибору платформи Аналітики в цьому столітті поставили безліч спеціалістів та організацій з інформатики та бізнес-аналітики. Тут немає єдиної відповіді для всього спектру вимог. Відповідь залишається дуже суб'єктивною, особливо враховуючи кількість ВІ та інструментів візуалізації на ринку. [11]

Для розгляду існуючих рішень Бізнес Аналітики було розглянуто дослідження Gartner. Ця консалтингова компанія випускає щорічний магічний квадрант для "Аналітики та платформ бізнес-аналітики" у лютому. Організації в різних областях та галузях звертаються до цього Магічного Квадранта, щоб зрозуміти, до якого інструменту потрібно звернути увагу, ось як це зазначає Gartner:

«Розширені можливості стають ключовими диференціаторами для аналітики та ВІ-платформ, в той час, коли хмарні екосистеми також впливають на рішення щодо відбору. Цей Магічний Квадрант допоможе лідерам даних та аналітики розвивати свої портфелі аналітики та технології ВІ в світлі цих змін. " – Gartner, 2020 [5].

Gartner використовує наведені нижче функції для ранжирування цих інструментів у "Чарівному квадранті":

- комплексна підтримка можливостей звітності підприємства;

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 22 |
| Зм | Лист | № документа | Підпис | Дата | | |

- розширена аналітика.



Рис 2.1. – Магічний квадрант Gartner 2020

Далі розглянуто декілька платформ з цього звіту.

3.1.1 Microsoft Power BI

Тринадцять років поспіль Gartner визнає Microsoft лідером магічних квадрантів у платформах аналітики та бізнес-аналітики. Він має всеосяжну та прозорливу карту розвитку продукту та масштабний ринковий доступ через свій канал Microsoft Office. [5]

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 23 |
| Зм | Лист | № документа | Підпис | Дата | | |

Дана платформа пропонує підготовку даних, засоби візуалізації даних, інтерактивні інформаційні панелі та розширену аналітику в Power BI. Він доступний як опція SaaS (Software as a Service), що працює в хмарі Azure, або як локальна опція на сервері звітів Power BI Report Server. Power BI Desktop можна використовувати як самостійний, безкоштовний особистий інструмент аналізу. Установка Power BI Desktop необхідна, коли користувачі опрацьовують складні масиви даних, що використовують локальні джерела даних. [11]

Microsoft випускає щотижневе оновлення своєї хмарної служби, яка додала сотні функцій у 2019 році. До останніх доповнень відносяться візуальні зображення декомпозиції, підключення даних соціальних мереж та вдосконалення географічного відображення.

Протягом років після запуску в 2013 році Power BI був продуктом «послідовником», який мав бути лише «досить хорошим», враховуючи ціну. Це вже не так – із випусками в 2019 році хмарний сервіс Power BI Pro обігнав більшість своїх конкурентів з точки зору функціональності. Це перевершило багатьох, включаючи інноваційні можливості для розширеної аналітики та автоматизованого управління машинного навчання. Служби, що працюють на штучному інтелекті, такі як текст, настрої та аналітика зображень, доступні в Power BI та користуються можливостями Azure. Переважна більшість опитаних респондентів Microsoft рекомендують Power BI користувачам без попередньої технічної підготовки.

Всеосяжність доступних засобів візуалізації: Microsoft продовжує інвестувати у широкий набір можливостей візуалізації та інтегрувати їх у Power BI.

У порівнянні з хмарною службою Power BI Pro, локалізовані пропозиції Microsoft мають значні функціональні прогалини, включаючи інформаційні панелі, аналітику потоку, попередньо вбудований вміст, природні мовні запитання, розширення (те, що Microsoft називає Quick Insights) та

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 24 |
| Зм | Лист | № документа | Підпис | Дата | | |

попередження. Жодна з цих функцій не підтримується на сервері звітів Power BI.

Power BI пропонує дуже широкий спектр засобів завантаження даних. На час останнього оновлення Power BI Report Server налічується 77 активних засобів підключення до різних джерел даних, серед яких: SQL Server Database, SQL Server Analysis Services, Azure SQL Database, Azure SQL Data Warehouse, Excel, Access Database, Active Directory, Amazon Redshift, JSON, MySQL Database, ODBC, OLE DB, та багато інших.

3.1.2 Tableau

Він пропонує засоби візуалізації даних та моделей даних, що дозволяє користувачам отримувати доступ, обробляти, аналізувати та представляти результати своїх даних.

У 2019 році Tableau значно розширив сферу своїх пропозицій, особливо їх розширену аналітику та можливості управління. Для розширеної аналітики Tableau представив як Ask Data, так і Explain Data, щоб забезпечити запит на природній мові та автоматизовану інформацію. Для управління Tableau вдосконалив Builder Prep Builder (який постачається разом із Creator Tableau) та запровадив програму Preap Conductor для планування та моніторингу завдань управління даними. Підготовчий провідник Tableau поставляється в комплекті з каталогом Tableau як частина надбудови управління даними. Tableau також представив надбудову Server Management, яка забезпечує управління сервером, міграцію вмісту та оптимізацію робочого навантаження. Tableau також перемістив значну частину своєї клієнтської бази в хмару за допомогою Tableau Online.

Однією зі значних переваг Tableau є простота візуалізації аналізу та маніпулювання даними. Tableau дозволяє користувачам швидко перебирати дані з широкого спектру джерел даних, поєднувати їх та візуалізувати

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІАз61.030БАК.005 ПЗ | Лист |
| | | | | | | 25 |
| Зм | Лист | № документа | Підпис | Дата | | |

результати, використовуючи кращі практики візуального сприйняття. Даними можна легко маніпулювати під час візуалізації, наприклад, під час створення груп, бункерів та ієрархій.

Незважаючи на нові випуски продуктів управління сервером, які додавали можливості управління та адміністрування у 2019 році, серед деяких референтних клієнтів Tableau зберігається уявлення про слабе управління та адміністрування.

Взагалі, компонент Tableau Viewer має необхідний функціонал для візуалізації даних, та є відносно недорогим. Однак із доповненням «Серверне управління» та «Управління даними» клієнти Tableau зіткнуться із високими цінами, а значить, вони повинні розраховувати на додаткову плату за новий функціонал.

3.1.3 Qlik

Провідне рішення ABI від Qlik, Qlik Sense, працює з унікальним асоціативним двигуном Qlik, який впродовж останніх 20 років працює на продуктах Qlik. Двигун дозволяє користувачам усіх рівнів кваліфікації поєднувати дані та вивчати інформацію без обмежень інструментів на основі запитів. Когнітивний механізм Qlik додає продукту функціональність штучного інтелекту і машинного навчання та працює з Associative Engine, щоб запропонувати методи для розширення аналізу.

Qlik продовжує вдосконалювати архітектуру на базі мікросервісів на платформі та можливості голосового вводу запитів. Повна версія SaaS Qlik Sense Enterprise доступна і є основою нового випробуваного досвіду на основі SaaS від Qlik. У червні 2019 року Qlik представив "асоціативні уявлення" як розширену здатність до аналітики, яка використовує когнітивний механізм Qlik для розкриття інакше прихованих уявлень

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 26 |
| Зм | Лист | № документа | Підпис | Дата | | |

Qlik був одним з перших постачальників, які запропонували можливість управління досвідом користувача. Гнучкість розгортання локальних послуг або будь-якого з найбільших постачальників хмарних технологій або використання комбінації обох підходів або використання повної пропозиції Salik від Qlik залишається в центрі зору Qlik.

Розширюваність можливостей платформи: портфоліо пропозицій Qlik охоплює ряд етапів життєвого циклу аналітики. Qlik Sense надає можливості для виявлення візуальних даних для самообслуговування для аналітиків та бізнес-користувачів, а також підтримує вбудовану розробником аналітику з тієї ж платформи. Qlik Data Catalyst використовується для каталогізації та додаткового управління. Крім того, хоча платформа інтеграції даних Qlik (раніше Attunity) є окремим програмним продуктом, вона додає потужні можливості інтеграції та руху даних.

Функція асоціативного розуміння використовує унікальний "асоціативний досвід" Qlik для автоматичного виявлення даних про дані, які в іншому випадку можуть бути пропущені інструментами на основі запитів. Проект Qlik Data Literacy допомагає користувачам усіх рівнів, клієнтам Qlik чи ні, краще розуміти та використовувати дані. [5]

Продукт налічує багато незалежних компонент, що надає системі гнучкості в налаштуванні та використанні. При цьому інтеграція з поширеними засобами збереження та обробки даних має багато недоліків, що автоматично призводить до необхідності встановлення повного комплексу засобів цього провідника. Як наслідок вартість рішення різко зростає. Також значно підвищуються вимоги до технічного рівня користувачів, що вже не відповідає вимогам про самообслуговування користувачем.

Найкраще функціональним та фінансовим вимогам відповідає платформа Microsoft Power BI Report Server. Дане рішення має широкий функціонал візуалізації, гнучкі засоби для управління даними та вбудовану інтеграцію з необхідними типами джерел даних.

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 27 |
| Зм | Лист | № документа | Підпис | Дата | | |

Рівень управління даними. Врахувавши розмір оброблюваного знімку даних що подається до аналізу було прийняте рішення використати підхід Вітрини даних для розробки рівня даних. Цей підхід забезпечує чистоту та якість даних, відповідність даних доменній області аналітики та широкий спектр засобів програмної реалізації.

3.2 Реляційні бази даних

Виходячи з вимог до платформи, встановлених у розділі 2 було прийняте рішення реалізувати рівень управління та збереження даних з допомогою підходу Вітрини Даних. Кращим технічним методом реалізації цього підходу будуть реляційні бази даних. Далі розглянуто два рішення засобів розгортки реляційних баз даних.

3.2.1 База даних SQL Server 2017

SQL Server 2017 включає безліч нових функцій, покращення та підвищення продуктивності Database Engine.

Нове покоління удосконалень обробки запитів, які адаптують стратегії оптимізації до умов виконання робочого навантаження додатків. Для цієї першої версії сімейства функцій адаптивної обробки запитів є наступні вдосконалення: адаптивне приєднання пакетного режиму, зворотний зв'язок із наданням пакетної оперативної пам'яті та перемежоване виконання для функцій, що оцінюються в таблиці. [3]

Автоматична настройка баз даних забезпечує розуміння потенційних проблем із виконанням запитів, рекомендує рішення та може автоматично виправити виявлені проблеми.

SQL Server Database Engine включає в себе Database Engine, основний сервіс для зберігання, обробки та захисту даних, реплікації, повнотекстовий

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 28 |
| Зм | Лист | № документа | Підпис | Дата | | |

пошук, інструменти для управління реляційними та XML-даними, в інтеграції аналітики баз даних та інтеграції PolyBase для доступу до Hadoop та інших неоднорідних джерел даних та сервера Служби якості даних (DQS).

Служби аналізу служб Analysis Services включають інструменти для створення та управління он-лайн аналітичною обробкою (OLAP) та додатками для обміну даними.

Служби звітування (SSRS) включають серверні та клієнтські компоненти для створення, управління та розгортання табличних, матричних, графічних звітів та звітів у вільній формі. Служби звітування – це також розширювана платформа, яку можна використовувати для розробки програм звітів.

Послуги інтеграції Інтеграційні сервіси (SSIS) – це набір графічних інструментів та програмованих об'єктів для переміщення, копіювання та перетворення даних (ETL). Він також включає компонент Служби якості даних (DQS) для інтеграційних служб. [15]

Основні послуги з обслуговування даних Основні служби даних (MDS) – це рішення SQL Server для управління базовими даними. MDS може бути налаштований для управління будь-яким доменом (продукти, клієнти, акаунти) і включає ієрархії, детальну безпеку, транзакції, версію даних та бізнес-правила, а також надбудову для Excel, яка може використовуватися для управління даними.

Послуги машинного навчання (In-Database) підтримують розподілені, масштабовані рішення машинного навчання з використанням корпоративних джерел даних. У SQL Server 2016 підтримувалася мова R. SQL Server 2017 підтримує R і Python.

Сервер машинного навчання (автономний) підтримує розгортання розподілених масштабованих рішень для машинного навчання на декількох платформах та використання безлічі джерел даних підприємства, включаючи Linux та Hadoop. У SQL Server 2016 підтримувалася мова R. SQL Server 2017 підтримує R і Python.

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 29 |
| Зм | Лист | № документа | Підпис | Дата | | |

Пов'язані програмні засоби для роботи з даними SQL Server повністю інтегровані між собою та нараховують широкий спектр додатків. Студія управління SQL Server – це інтегрована середовище для доступу, налаштування, управління, адміністрування та розробки компонентів SQL Server. Студія управління дозволяє розробникам та адміністраторам усіх рівнів кваліфікації використовувати SQL Server.

SQL Server Configuration Manager забезпечує базове управління конфігурацією для служб SQL Server, протоколів сервера, клієнтських протоколів та псевдонімів клієнтів.

Профілер SQL Server надає графічний інтерфейс користувача для моніторингу екземплярів двигуна баз даних або служб аналізу.

База даних поради з налаштування двигунів баз даних допомагає створити оптимальні набори індексів, індексованих подань та розділів.

Клієнт якості даних надає надзвичайно простий та інтуїтивно зрозумілий графічний інтерфейс для підключення до сервера DQS та проведення операцій з очищення даних. Це також дозволяє централізовано контролювати різні дії, що виконуються під час операції очищення даних. [9]

Інструменти даних SQL Server надають середовище для побудови рішень для компонентів Business Intelligence: служб аналізу, служб звітування та служб інтеграції. Інструменти даних SQL Server також включають "Проекти баз даних", що забезпечує інтегроване середовище для розробників баз даних, щоб виконати всі свої роботи з розробки баз даних для будь-якої платформи SQL Server у Visual Studio. Розробники баз даних можуть використовувати розширений браузер сервера у Visual Studio, щоб легко створювати або редагувати об'єкти та дані бази даних або виконувати запити.

Компоненти підключення Встановлює компоненти для зв'язку між клієнтами та серверами та мережеві бібліотеки для DB-бібліотеки, ODBC та OLE DB.

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 30 |
| Зм | Лист | № документа | Підпис | Дата | | |

3.2.2 Amazon Aurora

Amazon Aurora – це сервіс реляційних баз даних, який поєднує швидкість і доступність комерційних баз даних високого класу з простотою та економічністю баз даних з відкритим кодом. Сумісне з MySQL видання Aurora забезпечує до 5 разів пропускну здатність стандартного MySQL, що працює на одному і тому ж апаратному забезпеченні, і дозволяє існуючим програмам та інструментам MySQL запускатись, не вимагаючи змін. [2]

Тестування на стандартних орієнтирах, таких як SysBench, показало до 5-кратного збільшення продуктивності в порівнянні з запасом MySQL на аналогічному обладнанні. Amazon Aurora використовує різноманітні програмні та апаратні методи, щоб гарантувати, що двигун бази даних здатний повністю використовувати наявні обчислення, пам'ять та мережу. Операції вводу / виводу використовують методи розподілених систем, такі як кворум для покращення послідовності роботи.

Використовуючи API RDS Amazon або консолі управління AWS, є можливість масштабування обчислювальних та ресурсів пам'яті, що живлять ваше розгортання вгору або вниз. Операції масштабування обчислень зазвичай завершуються через кілька хвилин.

Amazon Aurora автоматично збільшуватиме обсяг бази даних у міру зростання потреб у вашій базі даних. Обсяг буде зростати з кроком від 10 Гб до максимум 64 ТБ. [2]

Збільшити пропускну спроможність для підтримки великих запитів додатків, створивши до 15 реплік бази даних Aurora. Репліки Amazon Aurora діляться тим самим базовим сховищем, що і вихідний екземпляр, знижуючи витрати і уникаючи необхідності виконувати записи у вузлах реплікацій. Це звільняє більше процесорної потужності для обслуговування запитів читання і скорочує час затримки реплік – часто до одноцифрових мілісекунд. Aurora

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 31 |
| Зм | Лист | № документа | Підпис | Дата | | |

забезпечує кінцеву точку зчитування, щоб програма могла з'єднуватися без необхідності відслідковувати репліки під час їх додавання та видалення.

Amazon Aurora Serverless – це конфігурація автоматичного масштабування за запитом Aurora, де база даних автоматично запускається, вимикається та масштабується, збільшуючи або зменшуючи потужність, залежно від потреб вашої програми. Aurora Serverless дозволяє запускати базу даних у хмарі, не управляючи жодними екземплярами бази даних.

Спеціальні кінцеві точки дозволяють розподіляти та завантажувати балансів навантаження на різні набори екземплярів бази даних.

Паралельний запит Amazon Aurora забезпечує швидші аналітичні запити щодо поточних даних. Це може прискорити запити до 2 порядків, зберігаючи при цьому високу пропускну здатність підтримки для основного навантаження.

Для глобально розповсюджених програм можливо використовувати глобальну базу даних, де одна база даних Aurora може охоплювати декілька регіонів AWS, щоб забезпечити швидке локальне зчитування та швидке відновлення після аварій. Глобальна база даних використовує реплікацію на основі сховища для тиражування бази даних у кількох регіонах AWS із типовою затримкою менше 1 секунди. Ви можете використовувати вторинний регіон як варіант резервного копіювання, якщо вам потрібно швидко відновитися після регіональної деградації або відключення. База даних у вторинному регіоні може бути підвищена до повноцінного читання / запису менше ніж за 1 хвилину.

Кожен фрагмент об'єму вашої бази даних розміром 10 Гб реплікується шістьма способами в трьох зонах доступності. Зберігання Amazon Aurora стійке до відмов, прозоро обробляючи втрати до двох копій даних, не впливаючи на доступність запису в базу даних і до трьох копій, не впливаючи на доступність для читання. Це дозволяє відновити базу даних у будь-яку секунду протягом періоду зберігання, до останніх п'яти хвилин. Резервні копії Amazon Aurora є автоматичними, поступовими та безперервними та не впливають на продуктивність бази даних.

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 32 |
| Зм | Лист | № документа | Підпис | Дата | | |

Знімки DB – це ініційовані користувачем резервні, що зберігаються в Amazon S3, які зберігатимуться, поки їх явно не видалено. Вони використовують автоматизовані додаткові знімки, щоб зменшити час та необхідне зберігання.

Backtrack дозволяє швидко переміщувати базу даних до попереднього моменту, не потребуючи відновлення даних із резервної копії. Це дозволяє швидко відновитись після помилок користувача, наприклад, випадання неправильної таблиці або видалення неправильного рядка. Увімкнувши Backtrack, Aurora зберігатиме записи даних протягом зазначеної тривалості Backtrack.

Amazon Aurora дозволяє зашифрувати ваші бази даних за допомогою створених та керованих ключів через службу керування ключами AWS (KMS). У екземплярі бази даних, який працює з шифруванням Amazon Aurora, дані, що зберігаються в спокої в базовому сховищі, шифруються, як і автоматизовані резервні копії, знімки та репліки в одному кластері. Amazon Aurora використовує SSL (AES-256) для захисту даних у дорозі.

Amazon Aurora надає показники Amazon CloudWatch для ваших екземплярів БД без додаткової плати. Ви можете використовувати консоль управління AWS для перегляду понад 20 ключових оперативних показників для екземплярів вашої бази даних, включаючи обчислення, пам'ять, сховище, пропускну здатність запитів, коефіцієнт звернення кешу та активні з'єднання. Крім того, ви можете використовувати Enhanced Monitoring для збору показників з екземпляра операційної системи, на якому працює ваша база даних.

Amazon Aurora поєднує в собі безпеку, продуктивність, високу доступність та довговічність підприємства з низькою вартістю та простотою використання MySQL. Це робить її хорошою міграцією при переміщенні робочих навантажень із дорогих комерційних баз даних до AWS. Можливості MySQL роблять оптимальну базу даних для широкого спектру навантажень

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІАз61.030БАК.005 ПЗ | Лист |
| | | | | | | 33 |
| Зм | Лист | № документа | Підпис | Дата | | |

бази даних, від простих транзакційних програм до складних навантажень OLTP та OLAP зі складними SQL та зберігаються процедурами. Стандартні засоби імпорту та експорту MySQL працюють з Amazon Aurora.

Aurora пропонує можливості машинного навчання безпосередньо з бази даних, що дозволяє додавати прогнози на основі ML в свої програми через звичну мову програмування SQL. Завдяки простій, оптимізованій та безпечній інтеграції між послугами машинного навчання Aurora та AWS, наявний широкий вибір алгоритмів ML, не будуючи спеціальних інтеграцій або переміщуючи дані. [2]

Aurora може працювати спільно з Amazon RDS Proxy, повністю керованим, високодоступним проксі-сервером бази даних, який робить додатки більш масштабованими, стійкішими до відмов бази даних та більш безпечними. RDS Proxy дозволяє програмам об'єднувати та обмінюватися з'єднаннями, встановленими з базою даних, покращуючи ефективність бази даних та масштабованість програми. Це скорочує час відмови шляхом автоматичного підключення до нового екземпляра бази даних, зберігаючи з'єднання програми. Це підвищує безпеку завдяки інтеграції з AWS IAM та AWS Secrets Manager.

3.3 Обчислювальна хмара Amazon Elastic (Amazon EC2)

Хмарні обчислення – це доставка ІТ-ресурсів за запитом через Інтернет із методологією ціноутворення «плати за те, що використовуєш». Замість того, щоб купувати, володіти та підтримувати фізичні центри обробки даних та сервери, надається доступ до технологічних служб, таких як обчислювальна потужність, сховище та бази даних, за необхідності від хмарного постачальника послуг, таких як Amazon Web Services (AWS).

Організації будь-якого типу, розміру та галузі використовують хмару для найрізноманітніших випадків використання, таких як резервне копіювання даних, відновлення аварій, електронна пошта, віртуальні настільні комп'ютери,

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 34 |
| Зм | Лист | № документа | Підпис | Дата | | |

розробка та тестування програмного забезпечення, аналітика великих даних та веб-додатки, орієнтовані на клієнтів. Наприклад, медичні компанії використовують хмару для розробки більш персоналізованих методів лікування пацієнтів. Компанії, що надають фінансові послуги, використовують хмару для розкриття та запобігання шахрайству в реальному часі. А виробники відеоігор використовують хмару, щоб доставляти онлайн-ігри мільйонам гравців у всьому світі.

Хмара надає легкий доступ до широкого спектру технологій, що в свою чергу значно зменшує час на впровадження нових системних та програмних засобів та будувати майже будь-які інфраструктурні рішення. [14]

Системні ресурси є легко масштабованими вгору або вниз, щоб миттєво зростати та зменшувати потенціал відповідно змінам навантаження на системи та новим вимогам.

За допомогою хмари можна за лічені хвилини розширитись до нових географічних регіонів та розгорнути доступ у всьому світі. Розміщення програм у географічній близькості до кінцевих користувачів зменшує затримку та покращує їхній досвід.

Три основні типи хмарних обчислень включають інфраструктуру як послугу (IaaS), платформу як послугу (PaaS) та програмне забезпечення як послугу (SaaS). Кожен тип хмарних обчислень забезпечує різні рівні контролю, гнучкості та управління, щоб ви могли вибрати правильний набір послуг для своїх потреб.

Нижче детально розглянуто три підходи до типів хмарних обчислень:

IaaS містить основні будівельні блоки для хмарних ІТ. Зазвичай він забезпечує доступ до мережесх функцій, комп'ютерів (віртуального або спеціального обладнання) та місця для зберігання даних. IaaS надає вам найвищий рівень гнучкості та управління управлінням вашими ІТ-ресурсами. Він найбільш схожий на існуючі ІТ-ресурси, з якими знайомі багато ІТ-підрозділів та розробників.

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 35 |
| Зм | Лист | № документа | Підпис | Дата | | |

Платформа як послуга (PaaS) знімає необхідність в управлінні базовою інфраструктурою (як правило, апаратними та операційними системами) та дозволяє зосередитись на розгортанні та управлінні вашими програмами. Це допомагає вам бути більш ефективними, оскільки вам не потрібно турбуватися про закупівлю ресурсів, планування потужностей, обслуговування програмного забезпечення, виправлення або будь-який інший недиференційований важкий підйом, який бере участь у запуску програми. [15]

Програмне забезпечення як послуга (SaaS) надає повний продукт, яким керує та підтримує постачальник послуг. У більшості випадків люди, які посилаються на SaaS, посилаються на додатки для кінцевих користувачів (наприклад, веб-електронна пошта). За допомогою SaaS не потрібно думати про те, як підтримується послуга або як керується базова інфраструктура. Потрібно лише подумати про те, як ви будете використовувати саме це програмне забезпечення.

Amazon Elastic (Amazon EC2) – це веб-сервіс, який забезпечує безпечну та налаштовувану обчислювальну потужність у хмарі. Він покликаний полегшити розробникам проектувати хмарні обчислення в веб. Простий інтерфейс веб-сервісу Amazon EC2 дозволяє отримати та налаштувати інфраструктуру з мінімальними затратами робочого часу та фінансів. Він надає повний контроль над обчислювальними ресурсами та дозволяє працювати в перевіреному обчислювальному середовищі Amazon. Amazon EC2 представляє справжнє віртуальне обчислювальне середовище, що дозволяє використовувати інтерфейси веб-служб для запуску серверних вузлів з різними операційними системами, завантаження їх у користувацьке середовище додатків, управління доступом до мережі та інше.

Amazon EC2 надає ряд потужних функцій для побудови масштабованих, стійких до відмов корпоративних інфраструктурних рішень.

Сервери Amazon EC2 (Elastic compute cloud) надають прямий доступ до процесора та пам'яті базового сервера. Ці екземпляри ідеально підходять для

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 36 |
| Зм | Лист | № документа | Підпис | Дата | | |

робочих навантажень, які потребують доступу до апаратних наборів функцій (таких як Intel® VT-x), або для додатків, які потребують запуску в невіртуалізованих середовищах для ліцензування або підтримки вимог. Сервери побудовані за системою Nitro – це колекція апаратних засобів для завантаження та захисту обчислювальних одиниць, створених AWS, які об'єднуються, щоб надійно забезпечити високоефективні мережеві ресурси та сховища для серверів EC2.

За допомогою єдиного виклику API, флот Amazon EC2 дозволяє надавати обчислювальну потужність для типів екземплярів EC2, зон доступності та моделей придбання, щоб оптимізувати масштаб, продуктивність та вартість кінцевого комплексу рішень. Ви можете вказати, скільки ресурсів за-запитом та точково потрібно запустити через флот EC2. Також є можливість масштабувати ємність серверів на основі ядер, екземплярів або пам'яті. Також продукт надає сервіс автоматичного масштабування Amazon EC2 для забезпечення автоматичного корегування обчислювальної спроможності для типів екземплярів EC2, зон доступності та варіантів придбання в одній групі автоматичного масштабування.

Обчислювальні одиниці GPU. Клієнти, які потребують потужної потужності з плаваючою точкою, отримують перевагу від обчислювальних потужностей GPU загального призначення з примірниками AWS, Amazon EC2 P3, що мають до 8 процесорних ядер NVIDIA® V100 Tensor Core. Сервери типу P3 ідеально підходять для машинного навчання, високоефективних обчислень, динаміки обчислювальної рідини, обчислювального фінансування, сейсмічного аналізу, молекулярного моделювання, геноміки та надання робочих навантажень.

При потребі дуже високої, низької затримки, випадкового доступу вводу / виводу даних, можуть скористатися високопродуктивними серверами вводу / виводу. Такі засоби типу Amazon EC2 можуть надати клієнтам показники вводу-виводу понад 3 млн. операцій в секунду. До них відносяться сервери типу

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 37 |
| Зм | Лист | № документа | Підпис | Дата | | |

I3 та I3en, що обладнані SSD-накопичувачами на основі енергонезалежної пам'ять Express (NVMe), і вони ідеально підходять для клієнтів, які працюють із дуже високими показниками баз даних NoSQL, транзакційними системами та навантаженнями Elastic Search. Високі екземпляри вводу / виводу також пропонують послідовну пропускну здатність диска до 16 Гб / с, що ідеально підходить для аналітичних навантажень.

При потребі дуже високої щільності зберігання на примірник, і високий послідовний введення / виведення для таких інтенсивних додатків, як сховище даних з масовою паралельною обробкою (MPP), розподілені обчислення MapReduce і Hadoop, а також обробки журналів даних в нагоді можуть стати EC2 щільного зберігання. Сервер щільного зберігання – це машина типу Amazon EC2, яка може надавати клієнтам послідовний введення-вивід швидкістю до 3,9 Гб / с і надавати до 48 ТБ вбудованого об'єму пам'яті на 24 дисках жорсткого диска, або пропонувати баланс із меншим обсягом пам'яті і пам'ять на vCPU з мережею на базі ENA для до 25 Гбіт / с пропускну здатності мережі в групі розміщення.

Функція Optimize CPU надає повний контроль над своїми примірниками Amazon EC2 на двох фронтах. По-перше, надається можливість вказати власну кількість vCPU під час запуску нових серверів, щоб заощадити на витратах на ліцензування на основі vCPU. По-друге, наявна опція відключити багатопотоковість для робочих навантажень, які добре працюють з однопотоковими процесорами, такими як деякі високопродуктивні комп'ютерні програми (HPC).

Різні навантаження Amazon EC2 можуть мати дуже різні вимоги до зберігання. Крім вбудованого сховища для серверів, також є Amazon Elastic Block Store (Amazon EBS) і Amazon Elastic File System (Amazon EFS), щоб відповідати іншим вимогам навантаження на хмарне зберігання.

Amazon EBS забезпечує стійкі, високодоступні, послідовні об'єми блоку зберігання з низькою затримкою для використання з екземплярами Amazon

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 38 |
| Зм | Лист | № документа | Підпис | Дата | | |

EC2. Кожен диск Amazon EBS автоматично копіюється в зоні доступності, щоб захистити від збоїв компонентів, пропонуючи високу доступність та довговічність. Він призначений для менеджерів додатків, яким потрібно налаштувати навантаження на потужність, продуктивність та вартість.

Amazon EFS забезпечує простий, масштабований, стійкий, повністю керований хмарний зберігання файлів для спільного доступу. Розроблений для високої доступності та довговічності в декількох зонах доступності, він забезпечує інтерфейс файлової системи зі стандартною семантикою доступу до файлової системи, автоматично зростає і скорочується, а також забезпечує менеджерам програм високу пропускну здатність і стабільно низькі затримки в петабайтовій шкалі.[2]

Amazon EC2 забезпечує можливість розміщення інфраструктури у кількох місцях. Місцеположення Amazon EC2 складається з регіонів та зон доступності. Зони доступності – це окремі місця, які розроблені для ізоляції від збоїв в інших зонах доступності та забезпечують недороге підключення до низької затримки мережі до інших зон доступності в тому ж регіоні. Запустивши екземпляри в окремих зонах доступності, ви можете захистити свої програми від виходу з ладу через помилку в регіоні. Регіони складаються з однієї або декількох зон доступності, географічно розповсюджені та будуть розташовані в окремих географічних районах чи країнах. Зобов'язання Угоди про рівень послуг Amazon EC2 – це 99,99% доступності для кожного регіону Amazon EC2.

Еластичні IP-адреси – це статичні IP-адреси, призначені для динамічних хмарних обчислень. Еластична IP-адреса пов'язана з вашим обліковим записом, а не конкретним екземпляром, і ви керуєте цією адресою, поки не вирішите явно її звільнити. Однак, на відміну від традиційних статичних IP-адрес, Elastic IP-адреси дозволяють вам замаскувати збої конкретного вузла або зони доступності, програмуючи перекомпоновані ваші загальнодоступні IP-адреси до будь-якого примірника вашого облікового запису. Замість того, щоб чекати на спеціаліста даних, щоб він переналаштував або замінив свій хост, або чекав,

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 39 |
| Зм | Лист | № документа | Підпис | Дата | | |

коли DNS поширяться на всіх ваших клієнтів, Amazon EC2 дає змогу вирішити проблеми з вашим сервером чи програмним забезпеченням, швидко переставивши свою еластичну IP-адресу на екземпляр заміни. Крім того, ви можете додатково налаштувати зворотний запис DNS будь-якої вашої Elastic IP-адреси, заповнивши цю форму.

Автоматичне масштабування Amazon EC2 дозволяє автоматично змінювати об'єм Amazon EC2 вгору або вниз відповідно до визначених користувачем умов. Завдяки автоматичному масштабуванню EC2 ви можете гарантувати, що кількість екземплярів Amazon EC2, які ви використовуєте, масштабується плавно вгору під час стрибків користувацьких запитів для підтримки продуктивності, а також зменшується автоматично під час затримок попиту, щоб мінімізувати витрати. Автоматичне масштабування EC2 особливо добре підходить для програм, які піддаються часовим, щоденним або щотижневим змінам в користуванні. Автоматичне масштабування EC2 увімкнено Amazon CloudWatch і доступне без додаткової плати за виплату Amazon CloudWatch.

Клієнти з комплексними обчислювальними робочими навантаженнями, такими як щільно з'єднані паралельні процеси, або з додатками, чутливими до продуктивності мережі, можуть досягти таких же високих обчислень та продуктивності мережі, які забезпечуються побудованою на замовлення інфраструктурою, отримуючи вигоду від еластичності, гнучкості та економічних переваг Amazon EC2. Вузли кластерного обчислення, GPU кластера та кластери з високою пам'яттю були спеціально розроблені для забезпечення високої продуктивності мережі та можуть бути запрограмовані програмно запускатись у кластери – дозволяючи програмам отримувати низькопродуктивні мережеві показники, необхідні для щільно з'єднаних вузлів до вузла спілкування. Вузли кластерів також забезпечують значно збільшену пропускну спроможність, завдяки чому вони добре підходять для клієнтських додатків, яким необхідно виконувати операції, що залежать від мережі.

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 40 |
| Зм | Лист | № документа | Підпис | Дата | | |

Розширена мережа дозволяє отримати значно більшу продуктивність пакету в секунду (PPS), менший тремтіння мережі та менші затримки. Ця функція використовує новий стек віртуалізації мережі, який забезпечує більш високу продуктивність вводу / виводу та нижчу ефективність процесора порівняно з традиційними реалізаціями. Щоб скористатися розширеною мережею, слід запустити HVM AMI у VPC та встановити відповідний драйвер.

Адаптер еластичної тканини (EFA) – це мережевий інтерфейс для серверів Amazon EC2, який дозволяє клієнтам запускати додатки HPC (швидке з'єднання для кластерів HPC), що вимагають високого рівня взаємозв'язку між примірниками, наприклад, обчислювальної динаміки рідини, моделювання погоди та моделювання резервуарів, за шкалою на AWS. Він використовує побудовану на замовлення техніку обходу операційної системи для підвищення продуктивності міжвідомчих комунікацій, що є критично важливим для масштабування програм HPC. За допомогою EFA додатки HPC, що використовують популярні технології HPC, такі як інтерфейс передачі повідомлень (MPI), можуть масштабуватися до тисяч ядер процесора. EFA підтримує стандартні API фабрики бібліотек, тому додатки, які використовують підтримувану бібліотеку MPI, можуть бути переміщені до AWS з невеликими або відсутніми модифікаціями.

EFA доступний як додаткова функція мереж EC2, яку можна ввімкнути для типів серверів C5n.18xl, P3dn.24xl та I3en.24xl.

Клієнти можуть приватно отримувати доступ до API Amazon EC2 із своєї віртуальної приватної хмари Amazon (VPC) або через AWS Direct Connect, не використовуючи загальнодоступні IP-адреси та не вимагаючи, щоб трафік проходив через Інтернет. AWS PrivateLink – це спеціально створена технологія, розроблена для того, щоб клієнти мали доступ до послуг Amazon з високою продуктивністю та високою доступністю, зберігаючи весь мережевий трафік у мережі AWS. Щоб використовувати Amazon EC2 з AWS PrivateLink, потрібно

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 41 |
| Зм | Лист | № документа | Підпис | Дата | | |

створити кінцеву точку для EC2 у своєму VPC. Будь-який трафік, призначений для цієї кінцевої точки, буде приватно направлений до служби EC2.

Образи Amazon Machine (AMI) заздалегідь налаштовані зображення серверів з постійно зростаючим списком операційних систем. Операційні системи, які зараз доступні для використання Amazon EC2, включають:

- Amazon Linux;
- Microsoft Windows Server;
- CentOS;
- Debian

Висновки до розділу

Розглянувши наявні системні та програмні рішення на ринку та маючи конкретизовані вимоги до системи було прийняте кінцеве рішення про структурні компоненти реалізації платформи Бізнес Аналітики.

Аналітичні системи було оцінено згідно двох підходів:

а) оцінка провідними консалтинговими компаніями на ринку IT послуг. Експертні висновки таких компаній є першим ресурсом для звуження кола вибору систем Бізнес Аналітики. Було визначено три лідери для подальшого розгляду на впровадження;

б) відповідність користувацьким вимогам. Саме вимоги кінцевого користувача і є ключовим фактором вибору системи. Правильна організація роботи та залучення користувачів в процес тестування продукту дозволяє швидше впровадити інформаційний засіб та якісніше оцінити результативність роботи.

Для реалізації користувацьких засобів Бізнес аналітики було обрано платформу Microsoft Power BI Report Server.

Як засіб впровадження рівня даних було обрано реляційну баз даних Microsoft SQL server. Дане рішення забезпечує гнучкість управління даними,

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 42 |
| Зм | Лист | № документа | Підпис | Дата | | |

рівнем ізоляції на рівні баз та таблиць, має відповідний функціонал побудови ETL процесу для наповнення вітрини даних інформацією та ін.

Для реалізації системного рівня було обрано Amazon Elastic Compute Cloud (EC2), що надає широкий спектр засобів розгортання та управління комплексом системних засобів. Також він надає можливість налаштування рівнів доступу, шифрування даних і масштабування системи в разі пікового навантаження.

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІАз61.030БАК.005 ПЗ | Лист |
| | | | | | | 43 |
| Зм | Лист | № документа | Підпис | Дата | | |

4 РЕАЛІЗАЦІЯ ПРОГРАМНО-АПАРАТНИХ ЗАСОБІВ БІЗНЕС АНАЛІТИКИ

4.1 Потоки даних

Маючи визначені вимоги до розроблюваної системи було визначено сутності, засоби взаємодії з системою та потоки даних. Нижчим рівнем визначення потоків даних є Діаграма контекстного рівня, яка приведена в додатку ДЗ. [6]

Центральною системою взаємодії є рішення Бізнес аналітики. Операторами схеми є Інженер даних, Аналітик та Користувач. Ролі розділені за принципом рівня взаємодії з даними.

Інженер даних – роль, яка характеризує людину або групу людей, що відповідають за обробку та завантаження даних в систему з допомогою програмних засобів масової обробки даних. Вихідним потоком даних є очищені та агреговані дані, що передаються в систему. Вхідний потік даних представляє собою готову аналітику, що побудована на базі цих даних.

Аналітик – представник з доменної області аналізу, в даному проекті це Фінансовий аналітик. Це людина або група людей, що мають бізнес знання з фінансів та можуть на основі очищених даних опрацювавши їх використовуючи фінансові формули та правила перетворити їх в інформацію. З контекстної діаграми бачимо, що аналітик передає в центральний процес правила та показники, що відповідають формулам показників та отримують Аналітику (значущу інформацію) як вхідний потік. Розглядаючи та аналізуючи інформацію, отриману з аналітичної платформи Аналітик робить висновки про відповідне проблемне питання, приймає рішення про нагальність будь-яких дій та може створювати нові показники ефективності для деталізації інформації. Також аналітик може підтвердити або спростувати актуальність та надійність

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 44 |
| Зм | Лист | № документа | Підпис | Дата | | |

даних використовуючи доменні знання та формувати нові технічні вимоги до оброблюваного набору даних для Інженера даних.

Користувач – будь-який працівник компанії, що має доступ до бізнес аналітики. Може переглядати існуючі звіти та робити висновки з представленої інформації. Дана категорія взаємодії виділена окремо, так як цей інтерфейс має набагато менше впливу на безпосередньо систему Бізнес Аналітики, але більш відповідальний за активне використання інформації, що надається системою. Мінімальний вплив на дані в системі надається через засоби користувацької взаємодії для кластеризації даних програмними засобами платформи Аналітики без впливу на набір даних Вітрини Даних.

4.2 Затвердження алгоритму реалізації системи.

Закінчивши дослідження наявних програмно-технічних засобів та визначившись з набором використовуваних рішень було визначено взаємозалежності в реалізації окремих компонент. Враховуючи всі деталі обраних засобів автоматизації процесу аналізу та візуалізації аналітики алгоритм розгортки технічної інфраструктури матиме наступну структуру:

- а) реалізація системного рівня;
- б) реалізація рівня даних (Вітрина даних та ETL сервер);
- в) встановлення платформи бізнес аналітики.

4.2.1 Реалізація системного рівня

Налаштування AWS аккаунту.

Робота проводиться в зоні доступності EU-WEST для забезпечення геолокаційної близькості до користувача згідно пункту 3.3. в розділі 3. Створення VPC та підмереж з діапазоном динамічних IP адрес. Реєстрація VPC в корпоративній мережі компанії необхідна для:

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 45 |
| Зм | Лист | № документа | Підпис | Дата | | |

- забезпечення доступу системи до інших ресурсів всередині корпоративної мережі компанії;
- резервування адрес для уникнення дублювання IP адрес іншими технічними відділами;

Далі було створено IAM ролі для налаштування прав доступу до технічних компонент системи. Налаштування ролей проведено засобами AWS. В консолі управління IAM створено наступні параметри:

- назва;
- пароль;
- присвоєно групу доступу.

В групі доступу налаштовано рівень доступу кожного IAM користувача що належить групі. Один користувач може входити до декількох груп.

Встановлено сервери використовуючи EC2 консоль.

Після відкриття акаунта підписки на веб-сервіси Amazon (AWS), обліковий запис AWS автоматично підписується на всі служби AWS, включаючи Amazon EC2.

AWS використовує криптографію відкритого ключа, щоб захистити інформацію про вхід на сервера.

В консолі Amazon EC2 в меню Ключі пари створено пару ключів.

Amazon EC2 пов'язує відкритий ключ із іменем, яке вказано як ім'я ключа.

Формат файлу збереження ключа вибрано *.pem. Цей формат збереження ключа можна використовувати з OpenSSH pem.

Створення групи безпеки. Групи безпеки виконують роль брандмауера для пов'язаних серверів, контролюючи як вхідний, так і вихідний трафік на рівні сервера. Додано правила до групи безпеки, які дозволять підключитися до свого сервера зі своєї IP-адреси за допомогою SSH. Також додано правила, які дозволяють вхідний та вихідний HTTP та HTTPS доступ з діапазону корпоративних IP адрес.

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 46 |
| Зм | Лист | № документа | Підпис | Дата | | |

Після створених правил та груп безпеки, шифрування ключів встановлено сервери для відповідних сервісів.

Для Вітрини даних необхідно 2 сервери:

Основна репліка:

- тип r5.4xlarge;
- EBS Volumes: диск ОС 500 ГБ, загального призначення, диски для збереження даних: три диски розміром по 1 ТБ, 10000 IOPS;
- AMI: Операційна система Windows Server Datacenter 2019 with SQL Server Enterprise;
- інші параметри такі як мережа, групи безпеки, IAM обрано з встановлених у попередніх пунктах;

Другорядна репліка:

- тип x1e.2xlarge;
- EBS Volumes: диск ОС 500 ГБ, загального призначення; диски для збереження даних: три диски розміром по 1 ТБ, 10000 IOPS;
- AMI: операційна система Windows Server Datacenter 2019 with SQL Server Enterprise;
- інші параметри такі як мережа, групи безпеки, IAM обрано з встановлених у попередніх пунктах.

Два сервери для розгортання системи Power BI Report Server:

- тип r5.4xlarge;
- EBS Volumes: диск ОС 500 ГБ, загального призначення, диски для збереження даних: три диски розміром по 600GB, 7000 IOPS;
- AMI: операційна система Windows Server Datacenter 2019 with SQL Server Enterprise;
- інші параметри такі як мережа, групи безпеки, IAM обрано з встановлених у попередніх пунктах.

Сервер для ETL:

- тип x1e.xlarge;

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 47 |
| Зм | Лист | № документа | Підпис | Дата | | |

- EBS Volumes: диск ОС 500 ГБ, загального призначення, диски для збереження даних: три диски розміром по 1 ТБ, 10000 IOPS;
- AMI: Операційна система Windows Server Datacenter 2019 with SQL Server Enterprise;
- інші параметри такі як мережа, групи безпеки, IAM обрано з встановлених у попередніх пунктах.

Для забезпечення високоякісного сервісу та швидкості реагування на кризові ситуації налаштовано моніторинг роботи серверів. Для повідомлень про статус навантаження на сервери використано засіб CloudWatch Agent. З допомогою CloudWatch налаштовано дашборди та повідомлення про такі параметри:

- навантаження на процесор %;
- мережева статистика вхід/вихід в МБ;
- кількість операцій читання/запису з кожного з дисків;
- час затримки відповіді сервера на запит читання/запису.

Готовий звіт для серверу ETL наведено на Рисунку 4.1.

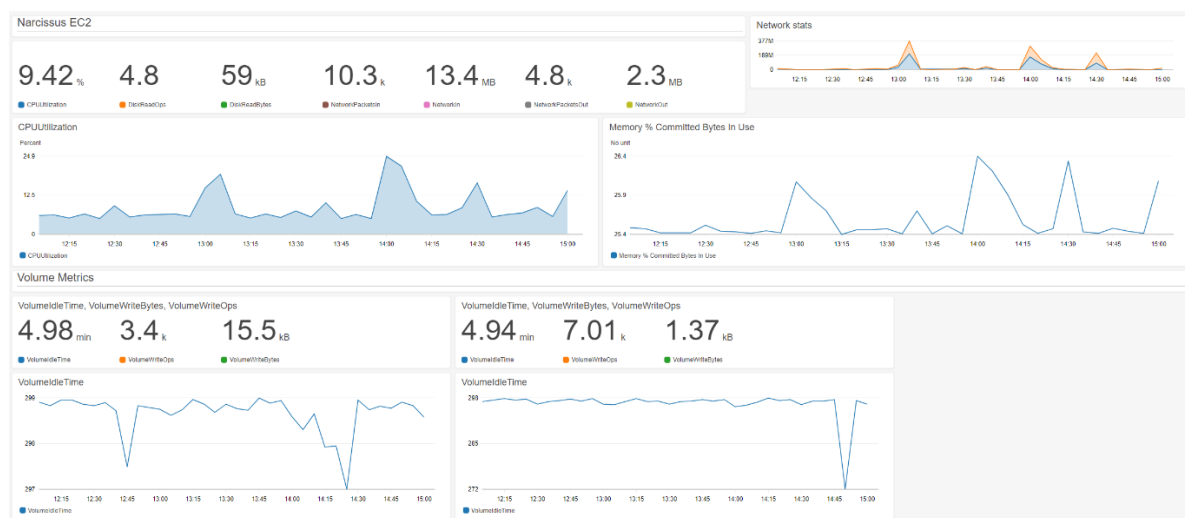


Рисунок 4.1. – CloudWatch звіт

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 48 |
| Зм | Лист | № документа | Підпис | Дата | | |

Після успішного завершення встановлення та налаштування всіх серверів засобами AWS було надано адміністративний рівень доступу групі адміністраторів технічної команди.

4.2.2 Налаштування Баз даних для Вітрини даних

Після завершення налаштувань серверів компоненти серверів баз даних Microsoft SQL Server також є доступними в корпоративній мережі, так як вони були включені в установку засобами AWS EC2. Скрипт бази даних Вітрини даних наведений у Додатку А.

При налаштуванні серверів баз даних пройдено наступні етапи.

Налаштовано права доступу до серверу бази даних – перед початком роботи з налаштуванням схем баз даних Вітрини даних треба впевнитися що команда адміністраторів даних має відповідні права для адміністрування сервера.

Створено безпосередньо Базу даних Вітрини Даних та відповідні таблиці. Таку ж базу даних розгорнуто на сервері реплікації. Таблиці спроектовані згідно наявних метаданих про оброблювану інформацію. Діаграма відношень сутностей бази даних додана до документації проекту в додатку Д2 [ІАз61.030БАК.005 Д2].

Налаштовано процес реплікації бази даних Вітрини Даних - знімки даних, зчитування журналів журналу та агент розповсюдження (push-підписка), створені на сервері розподілу. Завдання агента дистрибуції для підписок на передачу даних створено на абонентському сервері. Нове завдання дистрибуції відстежує стан (первинний або вторинний в групі доступу) баз даних розподілу та дозволяє або забороняє виконувати завдання реплікації на основі стану баз даних розподілу.

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІАз61.030БАК.005 ПЗ | Лист |
| | | | | | | 49 |
| Зм | Лист | № документа | Підпис | Дата | | |

Встановлено скрипти необхідні для підтримки серверів баз даних на всі наявні сервери. Скрипти створення окремих баз та створення збережених процедур для відслідковування дій користувачів та наведені в додатку Б.

4.2.3 Розгортання платформи Бізнес Аналітики

Рідний (Native) режим служби Reporting Services підтримує модель розгортання масштабу, яка дозволяє запускати кілька серверів звітності, які мають спільну базу даних одного сервера даних. Розгортання з допомогою декількох серверів було виконане для збільшення масштабованості серверів звітів для обробки великого навантаження конкурентних користувачів та великих завантажень обробки звітів та як задоволення необхідної умови для балансування завантаження декількох серверів звітів у кластері.

При налаштуванні системи сегментовано частини сервера звітів на різних серверах, використовуючи один сервер для інтерактивної обробки звітів та другий сервер для запланованої обробки звітів. У цьому випадку кожен екземпляр сервера обробляє різні типи запитів на один і той же вміст бази даних звітів, що зберігається в базі даних спільного сервера звітів.

При встановленні та налаштуванні платформи було виконано наступні дії. На двох серверах було виконано установку платформи засобами RsPowerBiReportServer.exe. Після встановлення програмного застосунку на основному сервері засобами Report Server Configuration Manager було виконано такі налаштування:

- Service Account: Virtual Service account;
- Virtual Directory: PBIRS_Reports;
- Database: Create new report server database;
- Database server: Serenity (local);
- Authentication type: Current User (integrated security);

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 50 |
| Зм | Лист | № документа | Підпис | Дата | | |

- Web Portal URL: IPv4, HTTPS Certificate, обрано встановлений сертифікат на сервері “serenity.ant.amazon.com”;
- Email sender address: serenity@amazon.com;

Після завершення налаштувань на основному сервері налаштування бази даних сервера виглядають як показано на Рисунок 4.2.

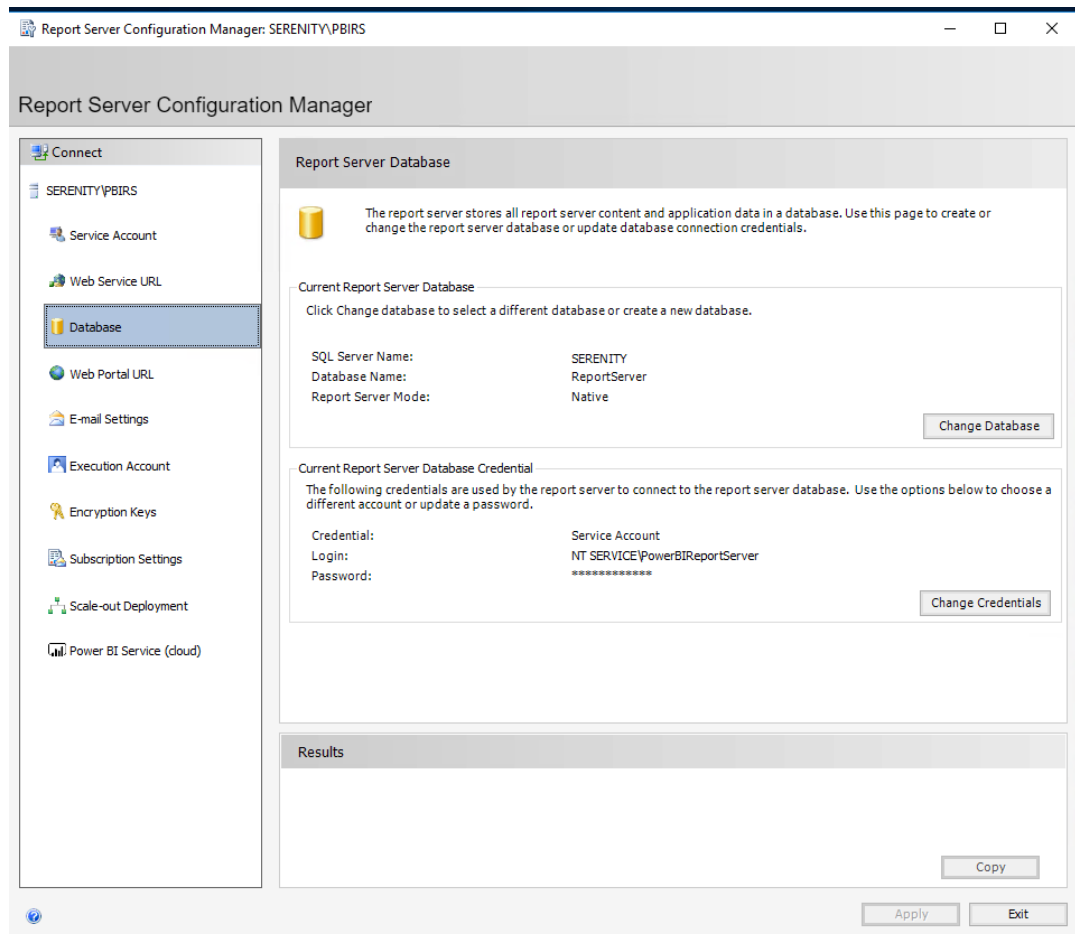


Рисунок 4.2. – Налаштування серверу звітів на основному сервері

На другорядному сервері, що використовується для виконання запланованих обробок звітів виконано наступні налаштування:

- Service Account: Virtual Service account;
- Virtual Directory: PBIRS_Reports;
- Database: Select Existing Report Server database;
- Database server: Serenity;
- Authentication type: Current User (integrated security);

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 51 |
| Зм | Лист | № документа | Підпис | Дата | | |

- Web Portal URL: IPv4, HTTPS Certificate, обрано встановлений сертифікат на сервері “serenity_as.ant.amazon.com”;
- Email sender address: serenity@amazon.com;

Після виконаних налаштувань підключення до серверу звітності на другорядному сервері виглядає як показано на рисунку 4.3.

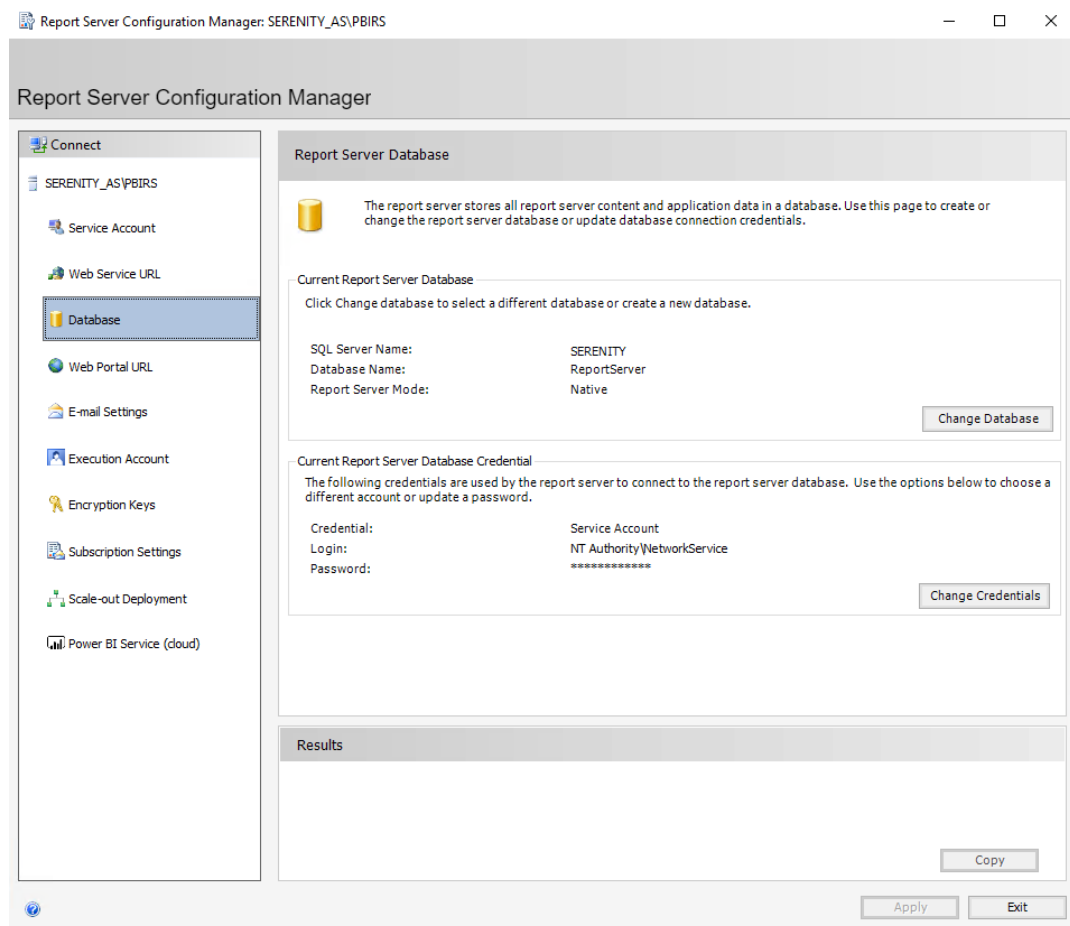


Рисунок 4.3. - Налаштування серверу звітів на другорядному сервері

Щоб приєднати другий екземпляр сервера звітів до розгортання масштабу було використано інструмент конфігурації служб звітування Перший сервер звітів вже ініціалізований для операцій з оборотним шифруванням, тому його можна використовувати для приєднання додаткових екземплярів сервера звітів до розгортання масштабу.

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 52 |
| Зм | Лист | № документа | Підпис | Дата | | |

На вкладці Розгортання масштабу з'явилося два записи, по одному для кожного екземпляра сервера звітів, підключеного до бази даних сервера звітів. Перший екземпляр сервера приєднаний. Другий сервер звітів знаходиться в статусі "Чекаю приєднання" як показано на Рисунок 4.4.

| Server | Instance | Status |
|----------|-----------------|-----------------|
| SERVER01 | MyNamedInstance | Waiting to Join |
| SERVER02 | MSSQLSERVER | Joined |

Рисунок 4.4. – Приєднання двох вузлів сервера звітності

Після під'єднання другорядного серверу для оновлень конфігурація має виглядає як показано на Рисунок 4.5.

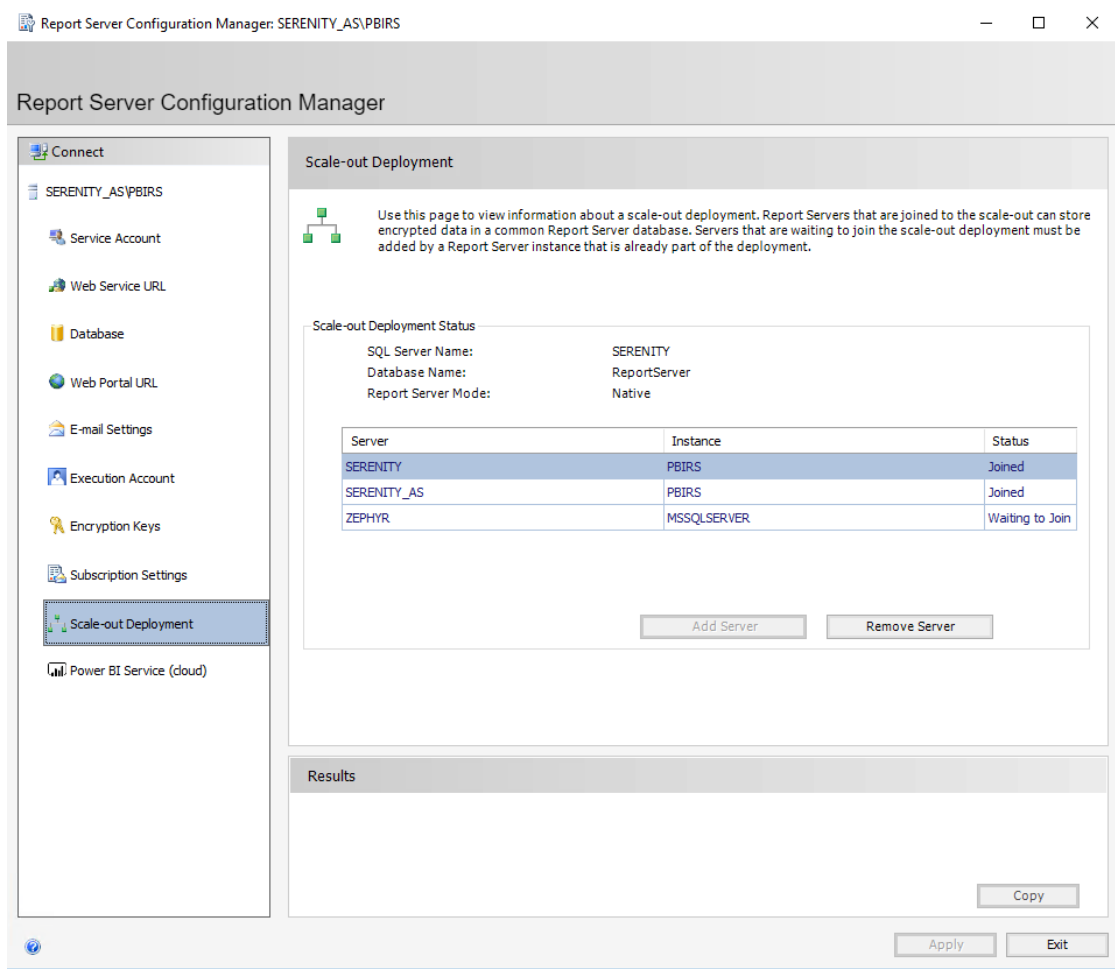


Рисунок 4.5. – конфігурація приєднаних серверів звітності

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 53 |
| Зм | Лист | № документа | Підпис | Дата | | |

Після завершення установки та налаштування засобів Серверів звітності та під'єднання серверів звітності виконано зміни в стандартних налаштуваннях служб. Згідно інфраструктурній схемі, що наведена в додатку Д1 бачимо, що основний сервер в розгортці використовується для відображення звітів та взаємодії з користувачами (BI Prod Server), при тому другорядний сервер в кластері використовується для обробки задач із запланованих оновлень моделей даних звітів. Для виконання цієї умови виконано зміни у файлах конфігурації RsReportServer.config основного та другорядного звітів у масштабі розгортки. На основному сервері змінено параметр IsDataModelRefreshService = False. Встановлення прапорця у значення False забезпечує вимкнення можливості виконання задач засобами лише основного сервера, значить все навантаження на оновлення звітів відправляється на другорядний сервер. При цьому звільнення другорядного сервера від виконання задач з відображення звітів та інтерактивного навантаження від користувачів та оптимізація самих оновлень виконана за допомогою зміни параметрів до таких величин:

 PollingInterval = 1 – швидкість зчитування черги на оновлення 1 секунда (стандартна величина 10 секунд);

 URLReservations – дана секція XML коду конфігурації була закоментована для вимкнення доступу до серверу звітності за URL адресою;

 IsAlertingService = False – встановлення цього прапорця забороняє відправку повідомлень про події сервера звітності з другорядного сервера;

 MaxQueueThreads = 50 – дане значення встановлене для обмеження кількості виконання задач з оновлення звітів в одиницю часу (величина за замовчуванням = 0, необмежена). Це дає змогу обмежити навантаження на сервер для забезпечення стабільності роботи модуля оновлення даних. При перевищенні цього значення всі оновлення, що мали бути виконані стають в чергу та при звільненні місця на виконання задачі оновлення моделі даних зчитуються за алгоритмом FIFO і подаються послідовно на виконання.

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 54 |
| Зм | Лист | № документа | Підпис | Дата | | |

Після завершення налаштувань основний та другорядний сервіси було перезавантажено для прийняття нових параметрів. Користувацький інтерфейс отриманої системи зображено на Рисунку 4.6.

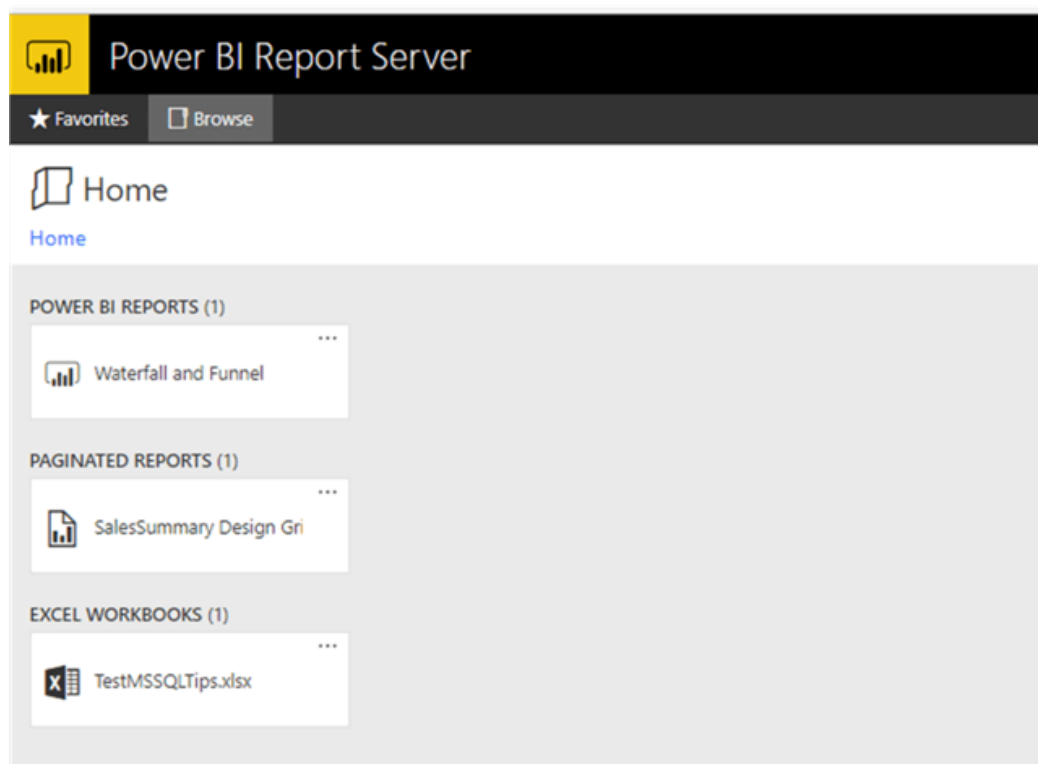


Рисунок 4.6 – Зображення користувацького інтерфейсу робочої платформи

4.3 Розробка Вітрини Даних

Згідно з наданими вимогами до домену та наповнення даних було визначено, що найкращий підхід для виконання завдання проекту з розробки рівня даних є Вітрина Даних. В Додатку Д2 наведено схему Бази даних.

Як бачимо з наведеної схеми База даних є нормалізована до Третьої Нормальної форми. Це дає змогу окремо розглянути два види матеріалів даних: факти та визначники. Виходячи з бізнес сфери в якій була поставлена задача говоримо, що Факти – це фінансові транзакції, що представляють вартісні

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 55 |
| Зм | Лист | № документа | Підпис | Дата | | |

показники витрат. Визначники – таблиці, що представляють словники, які додавши до таблиці Фактів дають змогу створити категорії фактів даних

Джерелом фактів даних є Озеро даних, що представлені в неагрегованому вигляді як фінансові транзакції.

Результуюча таблиця Вітрини Даних [dbo].[fct_Financials]. Дані оброблені програмним пакетом ETL агрегуються за наступними правилами:

- Date - group by
- ScenarioDimID - group by
- CompanyDimID - group by
- OpexNodeDimID - group by
- LocationDimID - group by
- CostCenterDimID - group by
- CostTypeDimID - group by
- AccountDimID - group by
- ChannelDimID - group by
- ProductDimID - group by
- Amount - sum
- SubmissionDimID - group by

Таблиці визначників заповнюються з системи Cognos та корпоративного сховища даних за допомогою ETL процесу реалізованого засобами SSIS. Процес завантаження та обробки даних налаштовано на окремому сервері для зняття навантаження з апаратних засобів серверів вітрини даних.

Відношення всіх визначників до таблиці фактів відповідає правилу «один до багатьох», що відповідає правилам нормалізації реляційних баз даних та забезпечує оптимальну швидкість операцій над групуванням та зчитування даних користувачем. Кожна таблиця визначника має властивість «ID», що є відповідним зовнішнім ключем при визначенні зв'язку з таблицею фактів.

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 56 |
| Зм | Лист | № документа | Підпис | Дата | | |

Перелік та опис таблиць визначників наведено нижче:

- ref.dim_Scenario – класифікація транзакцій за принципом фактична транзакція або план на наступний період
- ref.dim_Company – одиниця компанії. Зазвичай відповідає країні або економічному регіону
- ref.dim_CostCenter – класифікація за бухгалтерським принципом типу витрат
- ref.dim_Account – класифікація за бухгалтерським принципом типу транзакції
- ref.dim_OpexNode – тип операційних витрат
- ref.dim_Location – відповідна фізична одиниця для якої записана транзакція (склад, офіс, центр сортування, тощо)
- ref.dim_CostType – тип витрат
- ref.dim_Channel – відповідний канал доставки/відправки товару
- ref.dim_Product – характеристики продукту

Коди розгортання бази даних та створення таблиць наведено в програмних додатках до проекту.

4.4 Результати реалізації системи

Реалізована система пройшла комплексну перевірку та тестування за наступними критеріями:

- захист інформації;
- скорочення робочих затрат на побудову аналітики.

Захист інформації. Дані на жорстких дисках реалізованих у хмарному середовищі є зашифровані.

Всі паролі адміністратора захищені та зберігаються засобами AWS Odin.

Приватна мережа недоступна з мережі інтернет для випадкового клієнта

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 57 |
| Зм | Лист | № документа | Підпис | Дата | | |

SSL сертифікати встановлено на рівні веб клієнта для шифрування відображення інформації

Дані в Базах Даних Вітрини даних та Серверу звітів зашифровані програмно ключем кодування.

Відповідність вимогам клієнта. В ході реалізації було проведено тестування з обраною групою користувачів фінансового відділу на відповідність платформи раніше визначеним вимогам.

Програмні засоби мають доступний користувачеві функціонал що виконує функції:

- побудова та публікація звітів на веб клієнті платформи;
- обробка даних на виявлення закономірностей та підрахунок показників продуктивності;
- якість даним перевірена згідно підготованого набору даних та підтверджено, що система не видає помилок при опрацюванні даних;

Згідно результатам тестування систему було надано в користування відділу замовника.

Скорочення робочих затрат на побудови аналітики. Підраховано наступний вплив виражений в людино-годинах.

До впровадження системи на побудову звітності з підрахунку вартості на транспортування доставку товарів затрати на місяць складали 80 годин (1,5 робочі тижні одного аналітика), з яких:

- запит на отримання сирих даних з озера даних компанії: 4 години;
- підрахунки засобами EXCEL: 16 годин;
- періодичність: тиждень;

Після впровадження системи Бізнес аналітики затрати на побудову такого ж звіту складають 16 годин (2 робочі дні одного аналітика):

- побудова звіту в програмному засобі: 8 годин (одноразово);

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 58 |
| Зм | Лист | № документа | Підпис | Дата | | |

- перевірка якості даних та вірності підрахунків: 2 години (тиждень);

Витрати на розробку та підтримку даного звіту зменшились на 80%.

Висновки до розділу

Було виконано практичну реалізацію системи та оцінено її продуктивність.

Отримані результати тестування перевірки відповідності стандартам показують явний позитивний вплив на роботу відділу аналітиків:

- підвищення точності опрацювання даних через виключення фактору людської помилки
- зменшення затрат працівників на побудову звітів, що дає можливість професійним спеціалістам перейти на рівень підтримки прийняття рішень

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 59 |
| Зм | Лист | № документа | Підпис | Дата | | |

ВИСНОВКИ

За метою проекту розробки системи автоматизації фінансової звітності та аналітики було досліджено та побудовано комплекс засобів для автоматизації фінансової звітності.

Розглянуті методології підходу до зберігання та обробки даних що призначаються для аналітичного використання. В ході розгляду таких засобів було обрано підхід Вітрини даних – метод обробляє меншу кількість даних. Він забезпечує відповідність змісту даних предметній області дослідження та гнучкий в реалізації. Також менша кількість даних для обробки значно економить затрати на реалізацію цього рішення.

Для розгортання Бізнес Аналітики було обрано платформу Microsoft Power BI Report Server. Ця аналітична платформа показала себе як надзвичайно гнучка в користуванні аналітиками та включає в себе широкий спектр засобів візуалізації. Також наявність вбудованих засобів аналітичних обчислень вбудованих в користувацький інтерфейс надає можливість уніфікації джерела даних для його подальшого використання для побудови нової аналітики без потреби внесення змін в існуючі набори даних.

Проектування програмно-апаратних засобів для розгортки системи було виконано за методом згори-вниз.

Розробивши вимоги та визначившись з об'ємом опрацьовуваних даних, технічними вимогами до апаратних засобів що ставляться провідником аналітичних платформ та серверів баз даних було конкретизовано вимоги до серверної інфраструктури. Для реалізації цієї частини було обрано засоби AWS.

Дослідження застосунків хмарних обчислень підтвердило надзвичайну стабільність системи (групи доступності, резервне копіювання даних між регіонами), гнучкість засобів контролю показників продуктивності системних засобів та масштабованість (авто-розширювані групи серверів).

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 60 |
| Зм | Лист | № документа | Підпис | Дата | | |

В результаті тестування готового комплексу засобів було підтверджено відповідність системи поставленим вимогам та мету проекту досягнуто.

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 61 |
| Зм | Лист | № документа | Підпис | Дата | | |

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. В. Г. Халин [и др.] ; под редакцией В. Г. Халина, Г. В. Черновой. Системы поддержки принятия решений : учебник и практикум для бакалавриата и магистратуры — Москва : Издательство Юрайт, 2019. — 494 с.
2. AWS технічна документація по встановленню засобів хмарних обчислень, URL: <https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/dedicated-hosts-overview.html#dedicated-hosts-dedicated-instances>
3. MSDN Майкрософт, технічна документація з розгортки високопродуктивних реплікаційних вузлів, URL: <https://docs.microsoft.com/en-us/sql/database-engine/install-windows/install-sql-server-replication?view=sql-server-2016>
4. Уринцов А.И., Дик В.В., Системы поддержки принятия решений, М.: МЭСИ, 2008. – 205 с.
5. Аналітика Гартнер Магічний квадрант 2020, URL: https://www.gartner.com/doc/reprints?id=1-3TXXSLV&ct=170221&st=sb&ocid=mkto_eml_EM597235A1LA1
6. Илья Корнипаев. Требования для программного обеспечения: рекомендации по сбору и документированию — М.: Издательство «Книга по требованию», 2013 — 118 с.
7. Шаховська Н.Б., Пасічник В.В., Сховища та простори даних – Львів, Видавництво національного університету «Львівська політехніка», 2009 – 244 с.
8. В.М. Соловйов, Є.Я. Глушко, О.І. Олейніков, Я.В. Шрамко, В.І. Хорольський, О.А. Учитель, С.О. Семеріков, «Комп’ютерне моделювання та інформаційні технології в науці, економіці та освіті» – Кривий Ріг, Видавничий відділ КДПУ 2001, – Т. 2. – 299 с.

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 62 |
| Зм | Лист | № документа | Підпис | Дата | | |

9. Ставицький А. В. Навчально-методичний комплекс з курсів „Прогнозування” та „Фінансове прогнозування” / А.В. Ставицький. – К.: РВВ ІМФ, 2006. – 107 с.
10. Attansio, D. The multiple benefits of competitor intelligence / Attansio, D. // The Journal of Business Strategy, — 1998 — no. 3. — pp. 32-41
11. Методы и средства конкурентного анализа в бизнесе. / М.: БИНОМ. Лаборатория знаний, 2005, – 541 с
12. Himma K.E. The concept of information overload: A preliminary step in understanding the nature of a harmful information-related condition / K.E. Himma // Ethics and Information Technology. – 2007. – Vol. 9. – P. 259-272.
13. Амосов Н. М. Моделирование сложных систем. – К.: Наукова думка, 1968.
14. Бирман Э.Г. Сравнительный анализ методов прогнозирования / Э.Г. Бирман. – НТИ. Сер. 2. – 1986. – № 1. – С. 11–16.
15. Бююль, А. SPSS: искусство обработки информации. Анализ статистических данных и восстановление скрытых закономерностей : пер. с нем. – СПб. : ООО «ДиаСофтЮП», 2005. – 608 с.
16. Анісімов А.В., Кулябко П.П. Інформаційні системи та бази даних, Навчальний посібник для студентів факультету комп’ютерних наук та кібернетики. – Київ. – 2017. – 110 с.
17. Ульман Дж. Основы систем баз данных. М. Финансы и статистика. 1983.- 334 с.
18. Дрибас В.П. Реляционные модели баз данных. Изд-во БГУ. Минск. 1982.
19. Конноли Т., Бегг К., Страчан А. Базы данных: проектирование, реализация и сопровождение. Теория и практика. Издательский дом «Вильямс». 2000. – 1120 с.

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІАз61.030БАК.005 ПЗ | Лист |
| | | | | | | 63 |
| Зм | Лист | № документа | Підпис | Дата | | |

20. А. Ю. Берко, О. М. Верес, В. В. Пасічник, Системи баз даних та знань.
Книга 2. Системи управління базами даних та знань: навч. посібник. –
Львів : «Магнолія-2006». – 584 с.
21. Ситник Н. В. Проектування баз і сховищ даних: Навч. посібн. – К.:
КНЕУ, 2004. – 348 с.
22. Гайна Г.А. Основи проектування баз даних. Навч. посібник.-
К.:Кондор, 2007.- 208с.

| | | | | | | |
|----|------|-------------|--------|------|---------------------|------|
| | | | | | ІА361.030БАК.005 ПЗ | Лист |
| | | | | | | 64 |
| Зм | Лист | № документа | Підпис | Дата | | |